Redistribution with unobserved ex-ant choices

Carlos Eugênio E. da Costa

Setembro de 2003
Redistribution with Unobserved Ex-ante Choices.*

Carlos E. da Costa⁵
Getulio Vargas Graduate School of Economics

September 29, 2003

Abstract
Introducing dynamics to Mirrlees (1971) optimal taxation model creates a whole new set of issues that are only starting to be investigated in the literature. When choices are made before one’s realizing her productivity incentive constraints ought to be defined as a function of more complex strategies than in the static case. So far, all work has assumed that these choices are observable and can be contracted upon by the government. Here we investigate choices that: i) are not observed, and; ii) affect preferences conditional on the realization of types. In the simplest possible model where a non-trivial filtration is incorporated we show how these two characteristics make it necessary for IC constraints to be defined in terms of strategies rather than pure announcements. Tax prescriptions are derived, and it is shown that they bear some resemblance to classic optimal taxation results. We are able to show that in the most ‘natural’ cases return on capital ought to be taxed. However, we also show that the uniform taxation prescription of Atkinson and Stiglitz fails to hold, in general. Keywords: Optimal Taxation; Non-observability; Dynamic Contracts. JEL Classification: H21, D82.

1 Introduction
Despite its enormous impact, not only in the field of public economics but in all economic theory, the optimal income taxation framework of Mirrlees (1971) has lead to few fiscal policy prescriptions in all but one area: that of supplementary commodity taxation.

From the early contributions of Atkinson and Stiglitz (1976) and Mirrlees (1976) to the recent extensions of Naito (1999) and Saez (2001) this line of

*1 thank Luis Braido for helpful comments. The usual disclaimer applies.
⁵Address: 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: 55-21-2559.5495
research has permitted a better understanding of the role of commodity taxes when an optimally designed tax schedule is available.

In recent years the research agenda has incorporated, among other things, the discussion of intertemporal taxation. More precisely, the role of capital income taxation.

Introducing dynamics in Mirrlees’ framework, however, is not a simple task. The early attempts in this direction, notably da Costa and Werning (2001), use a series of assumptions that render the problem stationary, and allow for the mapping into an essentially static one.

More recently, however, Golosov et al. (2003) investigated a dynamic model with non-trivial evolution on a agents’ information sets. Incorporating non-trivial filtrations and, of course, requiring decisions to be adapted to them, is a rather complex matter, for the same type of mapping to a static problem used in da Costa and Werning (2000), is not possible in this case. Not surprisingly, a complete characterization of optimal tax structure is not feasible in Golosov et al. (2003). Nonetheless, the authors provide an important contribution to the area, by deriving sharp policy prescriptions for supplementary commodity taxation, now extended to include the discussion of capital income taxation. They show that, if preferences are separable between labor and leisure: i) an inverse Euler equation analogous to Rogerson’s (1985) result holds; ii) uniform taxation of goods is optimal as in Atkinson and Stiglitz (1976).

This type of assumption on the information structure is particularly important for life-cycle considerations, where many decisions - e.g. consumption early in life, number of children - are made before the full realization of one’s productivity. However, for this literature to arrive at more acceptable tax prescriptions the particular assumptions made on capital income taxation still demands closer scrutiny.

The point here is that a common feature of all these papers is the assumption - explicit or not - that decisions prior to the resolution of uncertainty are observable and can be 'contracted upon' by the planner. This means that tax rates on goods can be made fully non-linear across types and across time, along the lines of Atkinson and Stiglitz (1976) paper.

However, such a feature - marginal tax rates being type dependent - is particularly disturbing in the case of capital income taxation. The problem here is that tax rates on capital returns to implement the optimal contract ought to be: i) non-linear and ii) based on an agent’s type. This type being revealed only when return on capital is also realized.

---

1 The paper by Brito et al. (1990) consider an intertemporal optimal income taxation model along the lines of Mirrlees (1971). The paper deals with the use of non-stationary tax schedules as a way of mimicking randomization devices, and does not include supplementary commodity taxation.

2 Optimal taxation à la Mirrlees with changes in the information structure of the economy, however, appears in the literature earlier - Cremer and Gahvari (1995a, 1999) and da Costa and Werning (2000). These are two period models, with simplifying assumptions that allow for full characterization not available in Golosov and alii. (2003).

3 We shall be calling non-trivial a filtration $\mathcal{F} \equiv \{\mathcal{F}_t\}_{t=0}^T$, such that there is at least one $t \in (1,...,T)$ such that the $\sigma$-algebra $\mathcal{F}_t$ strictly contains $\mathcal{F}_{t-1}$. 

2
Linear tax rates on capital income are, however, pervasive. This is to be expected if there is a low cost for arbitraging on differences in marginal tax rates on capital income among agents. In fact, with low transactions costs and non-observability of trade by the planner, only anonymous (linear) taxes are feasible, which makes the derived contracts non-implementable.

In this paper we take the implementability issue as a serious restriction on the model economy and assume that marginal tax rates on capital can only depend on the information available at the moment the investment is made.

This simple restriction on the allowable set of instruments obliges us to change the way we approach the optimal taxation problem. The point is that three novelties of our model concur to make the optimal taxation problem very non-standard: i) evolving information sets; ii) non-observable 'ex-ante' choices, and; iii) choices affecting 'ex-post' preferences.

These three characteristics appear in various papers, but were never dealt with at the same time. The interesting novelty that comes up from their simultaneous appearance is that the set of 'ex-post' implementable allocations become endogenous. This requires not only the incentive compatibility (henceforth, IC) constraints to be defined in terms of announcement strategies rather than pure announcements, but also optimal off-equilibrium non-observable choices to be included.

We deal with these issues in our simple two type two periods model, similar to Stiglitz (1982). We are, then able to characterize the optimal system. In normal cases - that is, when the IC constraints bind in the usual direction - the no-taxation at the top result is still valid while the positive marginal tax rate for the least productive agent need not apply. Still in the normal case, savings ought to be taxed. This goes along with the findings of Cremer and Gahvari (1995, 1999) da Costa and Werning (2001) and Golosov et al. (2003).

However, the famous Atkinson and Stiglitz (1976) uniform taxation result, found in Golosov et al. (2003) breaks down in our setup - independently on which IC constraints bind. In fact we show that homotheticity is needed for uniform taxation to obtain.

It is interesting to relate our result to that found in Cremer et al. (2001b). They show how uniform taxation is usually not optimal, and how income effects become important, by introducing another dimension of unobserved heterogeneity: namely endowment of other goods. What is surprising in our result is that no other dimension of heterogeneity is introduced in the model explicitly.

We argue, nonetheless, that there is a subtle way in which heterogeneity appears in our framework. As we hope to make clear, agents off the equilibrium path have different 'endowments' from a second period perspective - pretty much like in Cremer et al. (2001b) - for they have different levels of savings. This is enough to make their result arise in a model with essentially no other source of heterogeneity but the one found in Mirrlees (1971).

---

4 See Guesnerie (1995) for an interesting exposition.
5 This is also necessary when all choices are observed if the filtration is not trivial.
6 See, for example, Stiglitz (1987).
The remainder of this paper is organized as follows. A general discussion of the literature and the way this paper relates to it is provided in section 1.1. The model economy is presented in section 2. Then, in section 2.1 the concept of equilibrium and the approach we adopt for tackling the problem is described. Optimal income taxation is characterized in section 3, with taxation on savings introduced in 4. Section 5 provides derivations of optimal linear taxation of commodities and discusses Atkinson and Stiglitz (1976) uniform taxation result. Extensions of the model and further discussions are provided in section 6. Section 7 concludes.

1.1 Related Literature

A recent trend in the agency literature, as applied to optimal taxation, has been the incorporation of a non-trivial filtration with the natural requirement that decisions be adapted to it.

Early incursions on this area are found in Cremer and Gahvari (1995, 1999) and da Costa and Werning (2000). These papers consider two period models, where agents make some decisions related to the consumption of pre-committed goods prior to their knowing their productivities: this is the simplest possible way in which a non-trivial filtration may be introduced. Optimal taxes are, then, derived but implementation is never discussed.

In a recent contribution to the area, Golosov et al. (2003), study a multi-period model with essentially no restrictions on the information filtration. Though a full characterization is not feasible, with such a degree of generality, they are able to derive some important tax prescriptions for the case of separable preferences.

A common feature of these papers is that they all assume that decisions prior to the resolution of uncertainty are observable and can be 'contracted upon' by the planner. This means that tax rates on goods can be made fully non-linear across types and across time. However, such a feature - marginal tax rates being type dependent - is particularly disturbing in the case of capital income taxation.

In fact, even though implementation is not discussed in Golosov et al. (2003), one thing is apparent from the proofs of their theorems: tax rates on capital income to implement the optimal contract are not measurable with respect to the information set available at the time when the investment is made. This is a consequence of the fact that capital income taxes needed to implement optimal allocations ought to be based on an agents realized type.

In two recent papers - Golosov et Tsyvinski (2003) and Kocherlakota (2003) - it is shown that implementation devices where capital income taxes are not based on the announcements for the period where types (and returns) are realized do not work, in general. Kocherlakota (2003) works with a linear tax on capital defined 'ex-post' and based on the agents announcement works as an implementation device", while Golosov et Tsyvinski (2001) show that a non-linear

\footnote{That is exactly where the tax scheme differs from the one proposed in Golosov et all}
tax on capital defined only in terms of information available prior to the an-
nouncement of type may implement the optimal allocations for a Diamond and 
Mirrlees (1978) disability insurance setup. This latter paper provides, therefore,
one exception to the general point, while the first shows that measurability with 
respect to a particular filtration is key in determining if the instrument imple-
ments the allocation.

There is at least one good reason why one should be suspicious about the 
assumption on the tax instruments used in these models. For them to be feasible 
arbitrage opportunities must be left in the market. Though two identical assets 
in our case a riskless asset - have different returns for different agents, they 
are not allowed to arbitrage on these differences by trading. Ruling out this 
form of trade may require implausibly high transaction costs, or a high degree 
of observability and control of trades from the government. This is hardly the 
case for financial markets where finding a counterpart interested in your product 
- money - is quite easy.

In this paper, we take a different path. Instead of ignoring these imple-
mentation issues, we drop the assumption of non-linear taxes on capital income 
and follow Cremer et al. (2001a) in considering that savings decisions are not 
observed by the planner. The government is still able to tax anonymous trans-
actions and create a wedge between lending and borrowing rates. However every 
agent must face the same tax rate on capital.

Because there is no aggregate uncertainty, this is equivalent to requiring the 
tax rate on capital income to be measurable with respect to period 0 information 
set. In fact, this is probably the most appealing feature of the restriction we 
impose on tax instruments - measurability of capital income taxes with respect 
to the information set available at the moment the investment takes place.

The simultaneous appearance of three issues, never considered in a same op-
timal taxation framework, to the best of our knowledge, results in the problem 
being quite non-standard. First, agents do not have all the information at the 
time when some of the decisions are made. That is, a non-trivial filtration is 
introduced in the model for the agent in what regards her productivity. Deci-
sions are, then, required to be adapted to this filtration, along the lines of 
Golosov et al. (2003). Second, savings are not observable, so we are restricted 
to anonymous taxes on capital income as in Cremer et al. (2001a). Unlike them, 
however, we add the third aspect we referred to. We let savings affect 'ex-post' 
preferences.

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

This is simple to understand. Once preferences are changed, the ranking of 
two bundles may be inverted, and what was incentive compatible for one specific 
'ex-ante' choice may not to be for another. Seen from a different perspective,
the possibility of anticipating the pattern of announcements, conditional on realized types, by agents, will make off-equilibrium (from the perspective of second period choices) savings crucial in defining the set of feasible allocations.

In discussing the role of commodity taxation these off-equilibrium savings are important in showing how the uniform tax prescription of Atkinson and Stiglitz (1976) breaks down. Because commodity taxes affect the marginal utility of income, even if preferences are additively separable, they do play a role in punishing deviant behavior. A similar type of result appears in Cremer et al. (2001b), where exogenous heterogeneity of endowments creates this wealth effect we only get off-equilibrium.

Finally we show how the model developed here has important new insights about optimal taxation in a life-cycle framework.

2 The Environment

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

\[ v(x) + E[u(c, l)]. \]

where \( x \) is first period consumption and \( c \) and \( l \) are, respectively, second period consumption and labor. Whilst in this section 6.1 we consider first period consumption and labor supply, for most of the discussions throughout the paper we shall think of first period labor as being supplied inelastically. This will simplify statements a great deal, and help isolate the novel issues brought in by our setup.

Function \( v \) is twice continuously differentiable, (strictly) increasing and concave. As for \( u \), we take it to be twice continuously differentiable in its two arguments. It is also (strictly) increasing in \( c \), decreasing in \( l \) and concave.

Later on we shall also assume that both leisure and second period consumption are normal goods. At this moment the assumptions above stated are all we need. Notice, however, that the structure we impose on \( v \) guarantees normality of \( x \).

Uncertainty arises in this problem because at the first period agents do not know their 'adult' productivities, which we shall call their types. Once uncertainty is realized, preferences are the same irrespective of one's type; productivity is the only dimension of heterogeneity 'ex-post'. An agent of productivity \( w \) needs \( l = Y/w \) hours to produce output \( Y \). Hence, the higher the productivity, the more leisure one agent gets for the same output she produces.

To make the problem as simple as possible we follow Stiglitz (1982) in considering only two possible types: \( H \) (for high productivity) and \( L \) (for low). We also assume that they are in equal proportion in the population for notational simplicity.

With this assumptions, (1) becomes

\[ v(x) + \frac{1}{2} \left[ u(c^H, l^H) + u(c^L, l^L) \right]. \]
We let shocks be independent across agents so that we may take the 'ex-post' distribution of types to coincide with the 'ex-ante' one.

The economy is divided in two sectors: a production and a consumption sectors. The production sector transforms labor inputs $lw$ in output $Y$ with a linear technology. Output may also be transferred inter-temporally using a linear technology. That is, the production sector transforms one unit of consumption today in one unit of consumption tomorrow. We also assume that each agent is endowed with a stock $I$ of first period consumption goods. This stock will be assumed to be 'large enough' in order to avoid the equilibrium level of $s$ to be negative. In this very simplified model, this assumption is just to guarantee that consumption does not take place before production.

Finally, there is a benevolent government who inhabits this economy and maximizes the agents’ expected utilities. One could think of the redistributive motives arising from a utilitarian social welfare function. At this point this specification is not so important because agents are 'ex-ante' identical, so what matters is that the social welfare function introduces neither some extra degree of risk aversion, nor a different degree of impatience to the problem.

Following the tradition founded by Mirrlees (1971), the set of instruments available for the government to pursue its goals is not imposed in an 'ad hoc' fashion but is derived from the informational structure defined for the problem. So we continue the description of the environment with its informational structure.

First, as is standard in optimal taxation, we assume that once uncertainty is realized each agent’s productivity is only observed by the agent herself. Though productivity is not directly observable output produced by each agent is observed by everyone, which makes it negotiable. It is assumed to be traded only at a prohibitively high cost outside the production sector, however.

Transactions between sectors are observable by the government but transactions within sectors are not. Therefore, because labor can only be traded - at reasonable cost - with the production sector, and because these transactions are observed by the government, the use of a non-linear income tax schedule is possible.

We assume, however that the trade of goods across time can be done between agents - outside the production sector - at no transaction costs. This rules out the possibility of having the government directly control savings.

Next we discuss in more details the taxation problem.

### 2.1 The Direct Mechanism

To find the optimal tax schedule we proceed as usual and define a truthful direct mechanism. Given the optimal allocations, we invoke the taxation principle to map it into an optimal tax schedule.

---

8One way to avoid introducing the assumption on $I$, would be to model an overlapping generations model. In this case, the young would trade with the 'middle-aged' (we don’t have the elderly, yet.) and there would be no problem of this kind. However this would just create too much distraction from the main points we wish to emphasize.
The problem here is non-standard, however, because some choices are made before the revelation of types. Hence, we dedicate the next few pages to discuss how to characterize the optimal allocation.

2.1.1 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 budget set \( \{(y^L, y^L), (y^H, y^H)\} \). In section 4 we will assume that, at this stage of the game, a tax rate on savings \( R - 1 \) is also defined by the government, while in section 5 linear taxation of all goods is permitted. For now, choosing that budget set is all the government does in \( t = 1 \).

The agent follows, with the next move, by deciding how much to consume (or, equivalently, how much to save). This decision is made before nature defines the agent’s type. Once uncertainty is realized the agent chooses her bundle from the budget set offered by the principal. That is the timing of the game is as follows

<table>
<thead>
<tr>
<th>( t = 0.5 )</th>
<th>( t = 1 )</th>
<th>( t = 1.5 )</th>
<th>( t = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>govt chooses</td>
<td>agents choose</td>
<td>nature chooses</td>
<td>agents pick</td>
</tr>
<tr>
<td>Budget set</td>
<td>savings</td>
<td>types</td>
<td>bundle</td>
</tr>
</tbody>
</table>

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

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

To understand this point, we start with the description of an agent’s choice considering her possible actions in the second period. Actions are simply announcements (not necessarily truthful) about one’s type. The maximization problem in the second period, is simply the choice of an announcement, contingent on one’s type, that maximizes her utility. Notice that she takes the level of savings \( s \) as given, in the second period, when choosing what type to announce.

A strategy, however, is a rule that associates to each type a specific action. In this case, an announcement. That is, if we introduce a player, called nature, the strategy of an agent defines a response for each move by nature in the form
of an announcement of one’s type\textsuperscript{9,10}.

The choice of an optimal strategy is not done in isolation. When deciding her strategy, at the first period, the agent is also choosing a level of savings. In fact, because each strategy defines a strictly concave savings problem, we associate to each announcement a unique optimal level of savings.

Hence, when making first period decisions, one is in fact choosing a pair: the best announcement strategy and the corresponding expected utility maximizing level of savings. Formally, let $\sigma^k(j)$ be the announcement made in strategy $k$ if one realizes type $j$. There are, in this case, four possible strategies:

- $\sigma^1(H) = H; \quad \sigma^1(L) = L$
- $\sigma^2(H) = L; \quad \sigma^2(L) = L$
- $\sigma^3(H) = H; \quad \sigma^3(L) = H$
- $\sigma^4(H) = L; \quad \sigma^4(L) = H$

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

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

If we let $y_i$ and $y_i^*$ be, respectively, the income available and the output produced by agent of type $i$ (for $i = H, L$) and $I$ the endowment of an agent in the first period of her life, that is, before she realizes her ‘adult’ productivity, then we define $s^*$ by

$$s^* = \arg \max_s \left\{ 2u(I - s) + u(y^H + s, \frac{y^H}{w^H}) + u(y^L + s, \frac{y^L}{w^L}) \right\},$$

and $s^k$ by the analogous expressions at the allocations corresponding to the prescriptions of each strategy.

At this point it should already be clear that it does not suffice to consider IC constraints at the equilibrium level of savings. We have to consider off-equilibrium savings choices because, though truthful announcement may be the

\textsuperscript{9}We are using the terms strategy and action in a rather informal way. In any traditional tax game, one defines a strategy for a type as a prescription for an action given a first move by the government in defining a budget set. We are using the terms as if the choice by the government were not part of the game. What we want to stress here is that when solving a traditional taxation problem, IC constraints are there to guarantee that the allocations chosen by the government are such that each type chooses to make a truthful announcement as her optimal action. Here, however, the IC constraints must guarantee that the optimal strategy (or recommendation of action conditional on the move by nature) for the agent is to choose a truthful announcement no matter what type she finds out herself to be.

\textsuperscript{10}Notice that we are not allowing for random announcements. We don’t know, as of this moment, whether randomization is useful, beyond what the traditional literature teaches us—e.g. Brito et al. (1995).
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.

One might still be wondering whether, even if a truthful announcement is
the best strategy when all pairs \((\sigma, s)\) are considered, there could be a choice
of savings that would render the allocation \((y^*, Y^*)\) non-implementable, in the
second period. That would certainly be the case. However, this choice of savings
would lead to a strictly lower expected utility for the agent than the truthful
announcement. Any equilibrium of this kind for our Stackelberg game would
not be subgame perfect. Hence we may just proceed as discussed and set up
our problem.

2.1.2 The Redundancy of Second Period IC constraints

In principle, to set up the program to be solved by the government we should
include not only all the IC constraints that guarantee that truthful announce­
ment is the chosen strategy but also the traditional second period constraints.
However, as we will see, only two first period IC constraints need to be consid­

To get to this point we start by proving the redundancy of second period IC constraints. To do that, we shall first introduce some notation that will allow
us to economize on space and improve aesthetically the paper. Let

\[
U \left( x^i, Y^i, w^i \right) \equiv u \left( x^i, \frac{Y^i}{w^i} \right).
\]

Then, we define

\[
U^* (i) = U \left( y^i + s^*, Y^i, w^i \right)
\]

\[
U^* (i|j) = U \left( y^i + s^*, Y^i, w^j \right)
\]

\[
v^* = v \left( 1 - s^* \right)
\]

Analogously, we adopt the notation

\[
U^k (i) = U \left( y^i + s^k, Y^i, w^i \right),
\]

\[
U^k (i|j) = U \left( y^i + s^k, Y^i, w^j \right)
\]

\[
v^k = v \left( 1 - s^k \right)
\]

for \(k = 2, 3,\) and 4. It should also be clear in this case, that \(T (Y^H) = Y^H - y^H\)
and \(T (Y^L) = Y^L - y^L\).

What we show in proposition 1 is that, if the first period constraints are
satisfied then, the second period ones

\[
U^* (H) \geq U^* (L|H)
\]

and

\[
U^* (L) \geq U^* (H|L)
\]

will also be.
Proposition 1 \textit{Truthful announcement is the best strategy, if and only if}

\begin{align*}
(6) & \quad 2v^* + U^* (H) + U^* (L) \geq 2v^2 + U^2 (L|H) + U^2 (L) \\
(7) & \quad 2v^* + U^* (H) + U^* (L) \geq 2v^3 + U^3 (H) + U^3 (H|L) \\
(8) & \quad 2v^* + U^* (H) + U^* (L) \geq 2v^4 + U^4 (L|H) + U^4 (H|L)
\end{align*}

Hence, what we have shown is that, if the first period constraints are satisfied, so are the second period ones. Also from the proof it is readily seen that the constraints are satisfied with strict inequalities. As a consequence, it is apparent that any tax schedule will be interim inefficient, in the sense that once the saving decision is made, agents would want the government to redesign the tax schedule.

Conversely, considering the case of (6), if we required

\begin{equation}
(9) \quad U^* (H) = U^* (L|H)
\end{equation}

to be binding, we would have,

\begin{equation*}
2v^* + U^* (H) + U^* (L) = 2v^* + U^* (L) + U^* (L|H)
\end{equation*}

But then, generically,

\begin{equation*}
2v^* + U^* (H) + U^* (L) < 2v^2 + U^2 (L) + U^2 (L|H).
\end{equation*}

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

This results are akin to the ones found in the repeated moral hazard literature\cite{11}. Once savings choices are made agents would be better off if the government could redesign the tax schedule using (9). The 'contract' - for that matter, any deterministic implementable contract - is not renegotiation-proof in the sense of Dewatripont (1988).

Before writing down the Lagrangian for this problem we introduce one more assumption that will guarantee that strategy 4 can be ruled out on the grounds that there are no allocations that justify adopting it and that still satisfy the other constraints.

\textbf{Assumption A}: Both second period consumption and leisure are normal goods.

We should also notice is that if we write $m(.) = -U_Y/U_x$ it is easy to see that the above assumption implies that

\[\frac{\partial m(.)}{\partial w} < 0, \text{ and } \frac{\partial m(.)}{\partial s} < 0.\]

The first consequence is single-crossing (or Spence-Mirrlees condition), which results from normality of consumption. Usually single-crossing - which would suffice for our results - is directly assumed, but normality of consumption guarantees single-crossing in this setup with unidimensional heterogeneity, and is a

\textsuperscript{11} See Chiappori et Alli (1995), for example.
more intuitive assumption on preferences. The second one is due to normality of leisure and will turn out to be very important to many of our proofs.

One last piece of notation. Let

\[ m_k(j|j) = \frac{U_y(y^i + s^k, Y^i/w^j)}{U_y(y^i + s^k, Y^i/w^j)} \]

denote the marginal rate of substitution for an agent of type \( j \) who announces to be of type \( i \), given that she follows strategy \( k \). We shall also use \( m_k(j|j) \) for a generic allocation, for type \( j \), at the level of savings corresponding to strategy \( k \) and \( m_k(j) = m_k(j|j) \).

We are down to three first period IC constraints. What the next proposition shows is that if there is a level of savings that, at the same time, makes strategy \( 4 \) optimal, and gives the same expected utility that the truthful announcement, then one of the other two constraints is violated.

Proposition 2 If a budget set is such that there exists \( s^4 \) such that

\[ 2v^* + U^*(H) + U^*(L) = 2v^4 + U^4(H|L) + U^4(L|H) \]

and strategy \( 4 \) is the best announcement strategy for \( s^4 \), then there must be a level of savings that makes either strategy \( 2 \) or strategy \( 3 \) better than truthful announcement.

This shows that, whenever the first two constraints are satisfied, constraint (8) is satisfied as a strict inequality. Hence, we may always leave it aside.

### 3 Optimal Taxation

In this section we do not consider the possibility of taxing savings. This will allow us to make clearer statements about the nature of taxes. We come back to this issue in section 4.

Bearing the results and notation of previous section in mind, we are finally in a position to spell out the government’s program in a (almost) standard way,

\[ \max_{\{y^t, Y^t\} = H, L} 2v(I - s^*) + u\left( y^H + s^*, \frac{Y^H}{w^H}\right) + u\left( y^L + s^*, \frac{Y^L}{w^L}\right) \]

subject to the incentive compatibility constraints (6), (7) and the resource constraint

\[ y^H + y^L \leq Y^H + Y^L \]

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

The Lagrangian for the problem is

\[ L = 2v^* + U^* (H) + U^* (L) - \lambda \left[ y^H + y^L - (Y^H + Y^L) \right] \]

\[ + \mu^2 \left[ [2v^* + U^* (H) + U^* (L)] - [2v^2 + U^2 (L) + U^2 (L|H)] \right] \]

\[ + \mu^3 \left[ [2v^* + U^* (H) + U^* (L)] - [2v^3 + U^3 (H) + U^3 (H|L)] \right] \]

Let us start with the first order conditions with respect to \( y^H \) and \( Y^H \). Note that the terms related to savings all vanish. The logic is straightforward: we don’t have to worry about this effects as a simple application of the envelope theorem.

From the first order conditions with respect to \( y^H \) and \( Y^H \) we obtain

\[ m^* (H) = \frac{2\lambda - \mu^2 [U^2 (H) + U^2 (H|L)]}{2\lambda + \mu^3 [U^3 (H) + U^3 (H|L)]}, \]

where we use \( F_k \) to denote the partial derivative of function \( F \) with respect to variable \( k \).

Notice that if \( \mu^2 \neq 0 \), we cannot guarantee that \( m^* (H) = 1 \). That is, if there is a binding level of savings that makes it optimal for the low type to deviate by announcing to be a high type, then, in general, the marginal tax rate on the high type will not be zero.

Similarly, for \( y^L \) and \( Y^L \),

\[ m^* (L) = \frac{\lambda + \mu^2 [U^2 (L) + U^2 (L|H)]}{\lambda + \mu^2 [U^2 (L) + U^2 (L|H)]}. \]

If both constraints are binding we would expect the two agents to be distorted. That is we would expect marginal tax rates to be different from 0 for both types. Let us consider, however, the cases where one of the constraints is slack.

3.0.3 Case \( \mu^3 = 0 \)

The case \( \mu^3 = 0 \) is when it is only the behavior of high type pretending to be a low type that is of concern to us. In this case, which we claim to be the most appealing one, equation (12) becomes \( m^* (H) = 1 \). That is, the result of no distortion at the top is restored.

As for the low type, condition (13) remains. This condition is analogous to the ones found in the literature. There is an important difference, however. The expression would have \( U^* (L|H) \) and \( U^* (L|H) \) in the traditional problem, respectively, in the numerator and the denominator of the right hand side of this expression. Instead we have \( U^* (L) + U^* (L|H) \) and \( U^* (L) + U^* (L|H) \).

The question now is whether we can still guarantee that the marginal tax rate ought to be positive. Define,

\[ \gamma = \theta (L) m^2 (L) + (1 - \theta (L)) m^2 (L|H) \]

\[ \beta = \frac{\mu [U^2 (L) + U^2 (L|H)]}{\lambda}, \]
where
\[ \theta(L) = \frac{U_g^2(L)}{U_g^2(L) + U_g^2(L|H)} \]

That is, \( \gamma \) is the weighted average of shadow prices of leisure for a true type \( L \) and a high type pretending to be a low type - at level of savings \( s^2 \) - where the weights are defined by the marginal utility of income of the corresponding agents.

With this notation,

\[ m^*(L) = \frac{1 + \beta \gamma}{1 + \beta}, \text{ or} \]

\[ T'(L) = \frac{\beta}{1 + \beta} (1 - \gamma) \]

where we use \( T' \) without apologies.\(^\text{12}\)

It is interesting to notice that (15) implies that marginal tax rate for the low type and the average marginal taxes of the two type must have the same sign.

Before examining our case we recall that, in the traditional income taxation problem, \( \gamma = m(H|L) \). As a consequence of the SM, at any point \( (y,Y) \in \mathbb{R^2} \),

\[ m(L|L) > m(H|H) \]

it is, then, trivial to see that (15) can only be satisfied for \( T'(L) > 0 \).

Here, however, we have to compare the marginal tax rate for the low type along the equilibrium path, \( 1 - m^*(L) \), with an average between the tax rate of a high type who pretends to be a low type and a true low type, both evaluated at an off-equilibrium level of savings, \( \theta(L) m^2(L) + (1 - \theta(L)) m^2(L|H) \).

There are, in this case, some clear problems in applying the traditional argument to this case here. First of all, \( m^*(L) \) and \( m^2(L|H) \) are not defined in the same point of the \( C \times Y \) space (even though they are in the same point of the \( y \times Y \) space), so the inequality above need not hold. In fact, normality of leisure along with the fact that, usually, \( s^2 > s^* \) imply that we should expect \( m^2(L|H) > m^*(L|H) \). The fact \( m^*(L) > m^*(L|H) \) does not guarantee that \( m^*(L) > m^2(L|H) \). That is, the shadow price of leisure need not be higher for the low type than for the high type that would mimic her and change savings accordingly. Moreover, still considering \( s^2 > s^* \), we have \( m^2(L) > m^*(L) \), which goes unambiguously against our presumption that \( m^*(L) > \gamma \).

The conclusion is that we are not able to sign the marginal tax rate for a low type.

### 3.0.4 Case \( \mu^2 = 0 \)

The case where \( \mu^2 = 0 \), and \( \mu^3 > 0 \) is analogous to the former case. Now, it is the low type who should not be distorted. The high type should not be taxed.

\(^{12}\) It is a well known result in optimal taxation with a finite set of types that the tax schedule is not differentiable at the points where we find agents that are 'envied' by others. Here, this need not be the case. The tax schedule can be \( C^1 \). See figure ...for an example.
at the margin in this case. The point is that one may increase revenue from the high type while preserving her utility and still relax the IC constraint.

4 Capital Income Taxation

When it comes to capital income taxation prescriptions, there are two sets of results that stand in the literature in recent years. First, since the work of Chamley (1986), Judd (1987), and Lucas (1990), among others, it is known that in a Ramsey framework zero tax rate on capital income is typically optimal; be it a steady state prescription or just a rule to be satisfied on average.

Second, repeated agency models along the lines of Rogerson (1985) and, more recently, Golosov et al. (2003) all find an inverse Euler equation rule that in practice represent taxing future consumption as compared to present.

As argued before, this latter literature makes strong assumptions about the set of instruments available for the government. We, however, concentrate on linear taxes on capital income, which among other things makes the results comparable to the ones found in the former literature.

The procedure we shall adopt to find the optimal tax rate for savings, $R - 1$, is similar to the one found in Mirrlees (1976) and da Costa and Werning (2001, 2002).

To find the optimal tax rate on capital income we start by redefining $c_H = y_H + R s$ and $c_L = y_L + R s$ and rewriting the objective function as:

$$\max_{c_H, y_H, y_L} 2v(I - s') + u \left( y_H + R s', \frac{y_H}{w_H} \right) + u \left( y_L + R s', \frac{y_L}{w_L} \right)$$

It is also important to notice the dependence of $s$ on the second period allocations. In fact, abusing notation slightly, we have $s^* = s^* (y_H, y_H, y_L, y_L, I, R)$, $s^2 = s^2 (y_L, y_H, I, R)$, and $s^3 = s^3 (y_L, y_H, I, R)$.

Similarly we rewrite the government’s budget constraint as:

$$Y_L - y_L + (Y_H - y_H) + 2(1 - R) s^* \geq 0.$$  \hfill (16)

We then use the envelope theorem to find the derivative of the program with respect to $R$. Provided that the solution is interior, making this derivative equal to 0 will give us necessary conditions of our maximization problem. In practice, this may be done in many different ways; we opt to take the derivative of the Lagrangian and evaluate it at the optimum.

Differentiating the Lagrangian to this problem with respect to $R$, and equating it to 0 (i.e. $\partial L / \partial R = 0$) yields:

$$(1 + \mu^2 + \mu^3) \left( U_y^* (H) + U_y^* (L) \right) s^* - 2\lambda \left( s^* - (1 - R) s_H^* \right)$$

$$\mu^2 \left[ U_y^2 (L) + U_y^2 (H|L) \right] s^2 + \mu^3 \left[ U_y^3 (H) + U_y^3 (H|L) \right] s^3 \hfill (17)$$

Since we are evaluating it at the optimum, we can make use of the first order conditions with respect to $y_H$, $y_L$,

$$U_y^* (H) \left( 1 + \mu^2 + \mu^3 \right) - \mu^3 \left[ U_y^3 (H) + U_y^3 (H|L) \right] = \lambda (1 - (1 - R) s_H^* u). \hfill (18)$$
and $y^L$,

\begin{equation}
U_y^L (L) \left(1 + \mu^2 + \mu^3\right) - \mu^2 [U_y^R (L) + U_y^R (L|H)] = \lambda (1 - (1 - R) s^*_y),
\end{equation}

to simplify (17).

Notice that, (18) and (19) have an extra term, when compared to the expressions for the case where there is no taxation on savings. This accounts for the additional impact that changes in the contract have on government's revenues.

After some algebraic manipulation, expression (17) may be rearranged to yield

\begin{equation}
(1 - R) \left[ s_R^* - \left(s_L^* + s_H^*\right) s^* \right] = \frac{\mu^2}{2\lambda} [U_y^R (L) + U_y^R (L|H)] \left(s^2 - s^*\right) + \\
\frac{\mu^3}{2\lambda} [U_y^R (H) + U_y^R (H|L)] \left(s^3 - s^*\right)
\end{equation}

The reader who is familiar with Mirrlees's (1976) paper should have noticed the analogy between this formula and his tax prescription in the mixed taxation part. The left hand side is the 'discouragement' of $s$, while the right hand side has two terms related to the alternative strategies the agent would follow in off-equilibrium paths. In identifying the sign of $1 - R$, we shall attempt to provide further intuition for (20).

The first step is to determine the sign of the term inside brackets in the left hand side of (20). It is a form of slutsky equation, so that it can be interpreted as compensated savings, where the compensation is done through an increase in income in each state of the world. This is important, because, even though we can always guarantee that compensating demand of second period consumption increases, while first period's decreases, it is the fact that $I$ is held constant that guarantees that it translates directly into higher savings. What matters for us is that the sign is always positive.

Hence, the sign of $1 - R$ will be simply the sign of the left hand side of (20).

The two terms within brackets are clearly positive. As for $s^2 - s^*$, and $s^3 - s^*$ they represent simply the difference between savings of an agent who has decide to abide by the rules, and one who has decided to mimic a different type.

One important thing to realize is that they are of different signs. This makes it even harder to sign the optimal tax on savings. We will proceed as in the previous section, and consider only the special case that we think is most likely, namely the one where $\mu^2 > 0$ and $\mu^3 = 0$.

4.0.5 Case $\mu^3 = 0$

In this case, the deviation is that of a high type claiming to be a low type. Equation (20) becomes

\begin{equation}
(1 - R) \left[ s_R^* - \left(s_L^* + s_H^*\right) s^* \right] = \frac{\mu^2}{2\lambda} [U_y^R (L) + U_y^R (L|H)] \left(s^2 - s^*\right).
\end{equation}

\footnote{See Equation (86), in Mirrlees (76)}

\footnote{The proof is rather trivial so we spare the reader from having to go through it.}
If the intended mimicker increases her savings as compared to an intended abider - $s^2 > s^*$ - then savings ought to be taxed, i.e., $R < 1$. If the contrary occurs savings should be subsidized.

This is akin to what happens in Mirrlees' paper, with an important difference, it is not the mimicker, in the sense of a high type that announces to be a low type who should be taxed, but an agent who, at the first period, has decided to do that 'if she finds out to be of a high type'. This is because the relevant incentive compatibility constraint is the first period one.

In general we would expect $s^2 > s^*$. This is because an agent who intends to announce type $L$ no matter what she finds out her productivity to be, will receive less income - $y^L < y^H$. The expected marginal utility of income will tend to be higher than if she intended to make a truthful announcement. Therefore, she will increase her savings.

In fact if leisure is normal, then we can prove the following proposition.

**Proposition 3** Normality of leisure guarantees that $s^2 > s^*$.

Hence, combining proposition 3 and equation (21) we have the result that with normal leisure, savings ought to be taxed, as stated in the next corollary.

**Corollary 4** If leisure is normal and $\mu^3 = 0$, then it is optimal to tax savings (or subsidize first period consumption).

Next is an example where this is exactly the case.

**Example 5** Assume $u(x, l) = (x - c(l))$ with $\nu', \nu'' > 0$. In this case, $U_y(L, H) = U_y(L) > U_y(H)$, $y^L < y^H$. Therefore, $s^2 > s^*$, which separates taxation of savings - $R < 1$ - optimal.

The general result - and this example, in particular - also formalize the intuition provided by da Costa and Werning (2000) for the subsidization of pre-committed goods in a setting where first period consumption was observed. It is also related to the finding of Rogerson (1985) that in an optimal full-commitment contract for a dynamic moral hazard problem, the agent saves less than she would, if allowed to do so.

## 5 Atkinson-Stiglitz

In this section we investigate whether the Atkinson-Stiglitz (1976) uniform taxation result is still valid in this setup. In a model with non-linear taxation of all goods, Cremer and Gahvari (1995, 1999) and Golosov et al. (2003) show that AS extends to a dynamic setting. Here, we derive optimal tax formulae in our setup, and show the importance of controlling savings.

Throughout this section we take preferences to be as in example 5. It is well known that it guarantees the optimality of uniform taxation in Atkinson and Stiglitz (1976). The same is true for the setup of all other papers referred to
in the previous paragraph. We shall also assume for now that taxes may differ across ages. We do not claim realism for this assumption. It is made only for simplicity, and because it allows us to consider two issues we find important to address: the case where leisure is one of the goods consumed in the first period and; the inclusion of a retirement period to the model. The first of this issues fulfills the promise we made in Section 2 to allow for elastic labor supply, while the second introduces an important policy dimension in our model.

For most of our goals at this section the chosen level of labor supply will be irrelevant - that is exactly what separability delivers, so we need some new notation. First define

$$v(p, I - s) = \max_{x} \left\{ u(x) \mid xp \leq I - s \right\}$$

with $x(p, I - s)$ as the corresponding Marshallian demand.

Analogously,

$$v(q, y_i + s) = \max_{x} \left\{ u(x) \mid xq \leq y_i + s \right\} \quad i = H, L.$$ 

with $x(q, y_i + s) (i = H, L)$. It is important also do define the savings functions

$$s^*(q, p, I, y^H, y^L) = \arg\max_s v(p, I - s) + v(q, y^H + s) + v(q, y^L + s)$$

for savings along the equilibrium path, and

$$s^2(q, p, I, y^L) = \arg\max_s v(p, I - s) + v(q, y^L + s)$$

for savings off-equilibrium. We shall also work with the simplified notation

$$v^* = v(p, I - s^*)$$
$$v^*(i) = v(q, y^i + s^*) \quad i = H, L.$$ 

and

$$v^2 = v(p, I - s^2)$$
$$v^2(L) = v(q, y^L + s^2).$$

Finally let $x^*, x^*(H)$, $x^*(L)$ and $x^2(L)$, denote Marshallian demands with obvious notation.

In Appendix B we setup the government’s program and derive the following formulae for optimal taxation of second period good $i$.

$$\frac{2uv^2(L)}{\lambda} [x^2(L) - x^{*2}(L)] = -2s_i((p - 1)x^*_y + (q - 1)(x^*_y(H) + x^*_y(L)) - (q - 1)(x^*_y(H) + x^*_y(L))$$

(22)
where $s_i = s_i - (s_y x^i(H) + s_y x^i(L))$ and $x_i^* (H) - x_i^* (L)$. 

The discouragement of consumption of good $i$ has two components: the direct effect on compensated demands in each state of the world, and the indirect effect, through savings.

We consider two levels of AS. First, the conditions for $p = 1$ and $q = 1$ are very stringent. In fact we need $x^1_i (L) - x^2_i (L)$ to be constant across agents. That is the demand for all goods must vary by the same amount with $s^* - s^2$. This is obviously satisfied in the case where $s^* = s^2$.

Notice that homotheticity implies that $x^i(H) = \omega^i y(L) + s^*$ and $x^j(H) = \omega^j y(L) + s^2$. Hence, 

$$(x^i(L) - x^j(L)) - (x^i(L) - x^j(L)) = (\omega^i - \omega^j) (s^* - s^2)$$

which means that even homotheticity will not suffice. Once again we need $s^* = s^2$. This is what we should expect, anyway. As we have shown in example 5, for separable preferences, first period consumption ought to be subsidized.

Let us require, instead, that $q = 1$ but $p = \rho I$. Then, equation (31) becomes

$$(23) \quad \frac{\mu q^2}{\lambda} \left[ x^i(L) - x^j(L) \right] = (\rho - 1) I x^0 \hat{s}_i$$

In this case, we get proposition 6.

**Proposition 6** There is a $\rho$ that satisfies equation (23) only if

$$\frac{x^i(L) - x^j(L)}{x^0(H) + x^0(L)}$$

is constant across goods.

This condition is not satisfied, in general. However, if preferences are homothetic, $x^i = \omega^i y$, (where $\omega^i$ is a constant share) and $x^j = \omega^j$. Expression above becomes

$$s^* - s^2$$

and we get the following corollary.

**Corollary 7** If preferences are homothetic second period goods should be uniformly taxed.

For first period goods, an expression analogous to (23) is derived in appendix B.2. That is, we show that,

$$(24) \quad \frac{\mu q^2}{\lambda} \left[ x^j - x^0 \right] = \left[ (p - 1) x^0 + (q - l) x^0 \frac{x^0(H) + x^0(L)}{2} \right] \hat{s}_i + (p - 1) \hat{s}_i$$

15 To see that this is in fact the discouragement of consumption of good $i$ just notice that the vector $x^i(H) - x^j(H) x^{**} (H)$ is the derivative of the compensated (holding state utility constant) demands for all goods with respect to price $q^j$, which is equivalent to the gradient of compensated demand for good $i$. 

19
It is not hard to see that separability will not deliver the result.

Considering our previous discussion, we want to find conditions under which $q = 1$ and $p = 0.1$. Because the last term is a compensated demand, it vanishes at $p = 0.1$. Then, the expression becomes

$$2\mu x_j^2 \{x^j - x^{j+1}\} + \lambda x_j^a \{2(\rho - 1) \hat{s}_j\} = 0$$

Now, it is clear that

$$-\frac{2\mu x_j^2}{\lambda (x_j^a)} x^j - x^{j+1} = \rho - 1$$

Though we may not rewrite the above expression exactly as the one in proposition 6, it is still true that the only way for there to exist a $\rho$ such that it is satisfied for all $j$ is to have the left hand side of the expression independent of $j$.

Moreover, if is easy to see that with homothetic preferences the uniform taxation result holds. This is stated in the following corollary.

**Corollary 8** If preferences are homothetic first period goods should be uniformly taxed.

Notice that, even though, we are working in a Mirrlees setup, uniform taxation is only optimal if the conditions for the result in a Ramsey problem are present.

Cremer et al. also show that AS breaks down in a model where agents have different (non-observed) endowments. This extra dimension of heterogeneity is not needed in our model. In fact agents are heterogeneous here in only one dimension, as in Atkinson and Stiglitz (1976) and Mirrlees (1976). It must also be said that it is not the linearity restriction on commodity taxation that drives the result, since uniform taxes are linear.

There is however one sense in which one may argue 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 might explain why we get about the same prescriptions in a model where, unlike in their case, no extra dimension of heterogeneity was ever introduced.

### 6 Extensions

This section extends the results previously found to deal with three different issues: first period elastic labor supply, retirement and pre-committed goods. Sections 6.1 and 6.2 are particularly important from a policy perspective in that they allow for writing the model as a well specified - though still simplistic - life-cycle model.

Section 6.3 provides a discussion on the literature on pre-committed goods, and shows how, so far, all paper have (implicitly) worked with very unrealistic implementation devices.
6.1 First period labor supply

Incorporating first period labor supply, is now a simple task. Just let first period preferences be \( u (c, l) \). The price vector for first period will be \( p = \rho (1, w) \), where \( \rho \) is a positive scalar and \( w \) is after tax wage. This will allow for separating the discussion of inter-temporal distortion from the discussion of labor/leisure distortions.

For simplicity assume productivity to be 1. The interesting question is whether first period labor supply is to be distorted at the optimum, that is whether \( w f l \).

Notice that, when savings are observed, separability is sufficient to prove that labor/leisure choice should not be distorted. Here, however, we need not only separability, but also homotheticity. This is a direct consequence of equation (24) and corollary 8.

The point here is that, though there is no direct benefit in terms of relaxing IC constraints from distorting first period labor supply, there is an indirect gain from its effects on savings.

6.2 Retirement

To make this a complete (though very stylized) life-cycle model, we introduce a third period, where labor supply is set to zero.\(^{16}\) To do this just assume that there are \( n \) goods, but define \( q \in \mathbb{R}^n \). Inter-temporal separability would just mean that

\[
U (i) = U_1 (i) + U_2 (i)
\]

with \( U_1, U_2 : \mathbb{R}^{n+1} \to \mathbb{R} \).

We shall however, consider the more restricted case where

\[
U_1 (i) = \nu (x) - \zeta (l) \quad \text{and} \quad U_2 (i) = \nu (x)
\]

with \( \zeta (0) \) set to 0 and \( \nu (\cdot) \) homothetic.

It is rather easy to verify that uniform taxation is not to be expected if only separability is assumed. Once again it is the effect of relative prices on the marginal utility of income that creates the possibility of using differential taxation to affect savings - and, indirectly, IC constraints.

Homotheticity, however, implies that all goods within a period are to be uniformly taxed. We shall assume it to separate the discussion of inter-temporal distortions from that of within period distortions. So, the real interesting question is whether retirement savings are to be distorted and to what direction.

The answer is that it is optimal not to distort relative prices of second and third period consumption. That is, investments made in the second period should have a return equal to the marginal rate of transformation for these goods. First period investments, however, should still be taxed.

\(^{16}\) Assume, for instance that \( w = 0 \) for all agents at the third period.
Just note that a type $i$ agent gets, in the second period, produces $Y_i$ and gets income income $y_i$. Given a fixed amount of savings $s_i$ she optimizes,

$$\max_x v(x) + \nu (y_i + s_i - x) - \zeta \left( Y_i / w_i \right)$$

Just by examination, one can see that an agent does not alter the optimal balance between periods as a function of the announcement. It is always optimal to smooth consumption.

From the governments perspective, differential taxation across these periods does not help separating types.

Some words on the implicit assumptions adopted here are due. In practice, what this would be prescribing is retirement accounts that had differential taxation. This implies that agents may differ in the tax rates on their investments. Under the strict assumptions on the use of all arbitrage opportunities this type of policy would not be feasible. Governments, however, do provide tax incentives for retirement savings, which seems to point out that some form of non-linearity across time is used. Notice however that tax rates are still measurable with respect to the information set available at the moment the investment is made.

### 6.3 Pre-committed Goods

One final comment is worth making regarding pre-committed goods as defined by Cremer and Gahvari (1995a, 1995b). These are goods whose level of consumption must be defined prior to the realization of uncertainty. That is prior to agents knowing their productivities.

In terms of our model, this implies a generalization of preferences to

$$W(x, x(H), t(H), x(L), t(L)) = \frac{1}{2} v(x, x(H), t(H)) + \frac{1}{2} v(x, x(L), t(L))$$

instead of (2).

The same discussion here applies in a more radical form. Different choices of $x$ affect relative preferences between 'second period' goods directly and not only through the savings channel, as is the case when preferences are inter-temporally separable.

Observability is crucial here\textsuperscript{17}. More importantly any implementation device to work (for the general case) must be able to condition commodity taxes on the announcements, even when 'uniform taxation' is optimal.

\textsuperscript{17} It is interesting to note, however, that in Cremer and Gahvari (both, 1995 and 1999), the authors refer to implementation with linear taxes. By the same argument used to the case of savings one can see that this is not possible in general. Though in equilibrium agents have the same marginal rates of substitution for first period goods (since agents are 'ex-ante' identical) the tax schedule must be such that deviations from the equilibrium choice of first period goods must be strongly punished, which is not the case with linear taxes.
7 Conclusion

This paper incorporates aspects of life-cycle choices that have been recently introduced (never at the same time) in the definition of an optimal redistribution problem. On one hand, we included the possibility that choices be made before agents know their productivities (as in da Costa and Werning (2001) and Golosov et al. (2003)). On the other, we dropped the assumption that these choices were observable by the government (as in Cremer et al. (2001)). This latter aspect become very important because we allow for these choices to alter 'ex-post' preferences and, as a consequence, to break down traditional incentive compatibility constraints.

Implementation requires a different set of incentive constraints, which we impose in order to derive policy prescriptions. For the income tax schedule, we characterize tax systems and are able to identify what are the issues at stake in determining tax rates. Furthermore, in 'normal' cases, we show that marginal income tax should be 0 for the most productive agent. However we cannot guarantee that the least productive will face a non-positive marginal tax rate at the optimum.

Still in the normal cases, taxation of savings (or subsidization of early consumption) is shown to be optimal. Notice that this result applies only to investments made prior to the revelation of one's type. Hence, if we think of a simple life-cycle model where the third period is retirement, the return on investments made at this point in life ought not to be taxed - this feature goes along, in some degree, with actual tax systems.

From a purely theoretical perspective, our results are akin to the ones found in Golosov et al. (2003) and da Costa and Werning (2000), though we consider linear taxes on capital return, only.

Even sharper results are available for supplementary commodity taxation. We show, for instance, that this form of non-observability generates a violation of Atkinson and Stiglitz (1976) uniform taxation prescription. Homotheticity is also needed for the result to be valid. A similar result is obtained in Cremer et al. (2001), where another dimension of heterogeneity, in the form of unknown endowments is introduced.

What is surprising here is that we do not have another dimension of heterogeneity. At least not along the equilibrium path. The subtlety here is that agents do differ in another dimension - consumption endowments, as in Cremer et al. (2001) - from their off-equilibrium counterparts. This is enough for us to obtain the type of prescriptions they generate in a model with exogenous heterogeneity.

The model is admittedly simplistic and, though simplicity is important for highlighting the key issues at stake, it would be very interesting to explore in further detail the robustness of some of the results we found. However, this is no easy task, considering the non-trivial conceptual problems that arise in such framework.
References


A Proofs

Proof of Proposition 1. If one of these constraints is violated, there is a pair strategy/savings that is better than truthfully reporting one's type and choosing the corresponding maximizing level of savings. Hence, necessity. As for sufficiency, because savings are optimally chosen in (6) the right hand side
of this equation is greater than the same expression evaluated at any level of savings. In particular at \( s = s^* \), that is,

\[
2v^* + U^*(H) + U^*(L) > 2v^* + U^*(L|H) + U^*(L)
\]

\[
\therefore U^*(H) > U^*(L|H)
\]

And constraint (4) is satisfied with strict inequality. By the same argument, (7) implies that

\[
2v^* + U^*(H) + U^*(L) > 2v^* + U^*(H) + U^*(H|L)
\]

\[
\therefore U^*(L) > U^*(H|L)
\]

Which shows that (5) is also satisfied strictly. 

In order to prove proposition 2 we shall use the following fact, which generalizes a well known result to our setup.

**Claim 9** *Monotonicity is necessary for an allocation to be implementable*

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

\[
Z^H_+(y, Y) = \{(y', Y') \in \mathbb{R}^2; (x', y') \succ_H (y, Y)\}
\]

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

Similarly,

\[
Z^L_+(y, Y) = \{(y', Y') \in \mathbb{R}^2; (x', y') \succ_L (y, Y)\}
\]

Define also

\[
Z^H_-(y, Y) = \{(y', Y') \in \mathbb{R}^2; (x', y') \prec (y, Y)\}
\]

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

\[
(25)\quad Z^L_+(y, Y) \cap Z^H_+(y, Y) \subset Z^H_+(y, Y) \cap Z^L_+(y, Y)
\]

In fact, without loss, take a path starting at \((y, Y)\) such that

\[
dV(H) = V_y (|H|) [dy - m (|H|) dY] \geq 0 \Rightarrow
dy = m (|H|) dY
\]

In this case,

\[
dV(L) = V_y (|L|) [m (|H|) - m (|L|)] dY < 0
\]

for \(dY > 0\). If we define \(Z^- (y, Y) = \{(y', Y') \in \mathbb{R}^2; (x', y') < (y, Y)\}\), analogously, the same procedure shows that,

\[
(26)\quad Z^H_-(y, Y) \cap Z^-_-(y, Y) \subset Z^L_-(y, Y) \cap Z^-_-(y, Y)
\]

Now take a pair of allocations \((y^L, Y^L)\) and \((y^H, Y^H)\) such that \((y^H, Y^H) < (y^L, Y^L)\). From, (25) and (26), either \((y^L, Y^L) \in Z^H_+(y, Y)\), or \((y^H, Y^H) \in Z^L_+(y, Y)\). In either case, the allocation is not implementable. 

\[
\therefore
\]
We are now in a position to provide a simple proof of proposition 2.

**Proof of Proposition 2.** Assume that constraint (8) is binding at the optimal allocation. That is, there is a level of savings \( s^* \) such that
\[
2v^* + U^*(H) + U^*(L) = 2v^4 + U^4(H|L) + U^4(L|H).
\]
Notice that, because this is the optimal strategy, it must be the case that
\[
U^4(H|L) \geq U^4(L) \quad \text{and} \quad U^4(L|H) \geq U^4(H).
\]
From (25) and (26) it is rather straightforward to see that this requires \((y^H, y^H) < (y^L, y^L)\). But in this case, claim 9 implies that the allocation is not implementable.

**Proof of Proposition 3.** Let
\[
Z^+ (y^H, y^H) = \{(y', y') \in R^2; (x', y') \succ_H (y^H, y^H)\},
\]
at the equilibrium level of savings, \( s^* \), and
\[
Z^{-2} (y^H, y^H) = \{(y', y') \in R^2; (x', y') \succ^{-2} (y^H, y^H)\},
\]
at the optimal level of savings for strategy 2, \( s^2 \).

Because the allocation is implementable we know form claim 9 that \((y^L, y^L) \in Z_-(y^H, y^H)\) and \((y^L, y^L) \notin Z^+ (y^H, y^H)\). On the other hand, from the definition of strategy 2 it must be the case that \((y^L, y^L) \in Z^{-2} (y^H, y^H)\). However, if \( s^2 < s^* \), normality of leisure implies that
\[
Z^+ (y^H, y^H) \cap Z^{-2} (y^H, y^H) \subset Z^+ (y^H, y^H) \cap Z_-(y^H, y^H).
\]
A contradiction.

**Proof of Proposition 6.** First notice that
\[
\kappa s_i = v^y_y (H) x^i_y (H) + v^* y (H) x^i_y (H) + v^y_y (L) x^i_y (L) + v^* y (L) x^i_y (L)
\]
where \( \kappa \) is a negative constant that corresponds to the second order condition for the agent’s savings problem. Also notice that
\[
\kappa s_y = -V_{yy} (H) \quad \text{and} \quad \kappa s_y = -V_{yy} (L)
\]
Hence, from the definition of \( \delta_i \) we have that
\[
\kappa \delta_i = -2v^y_y \left[ x^i_y (H) + x^i_y (L) \right]
\]
But, then,
\[
(27) \quad \rho - 1 = \Phi \frac{x^{i*} (L) - x^{i2} (L)}{x^i_y (H) + x^i_y (L)} \quad \forall i
\]
with
\[
\Phi = -\frac{v^* y}{v^y y \lambda (1x^y_y)}
\]
Now just note that (27) can only be satisfied for all \( i \) if the right hand side is independent of \( i \).
B Derivation of AS formulae

The maximization problem for the government is

$$\max_{\{y', \mathcal{Y}\}_{i=n, L, P, q}} 2v^* + v^* (H) - \xi \left( \frac{Y^H}{w_H} \right) + v^* (L) - \xi \left( \frac{Y^L}{w_L} \right)$$

subject to

$$2v^* + v^* (H) + v^* (L) - \xi \left( \frac{Y^L}{w_L} \right) - \xi \left( \frac{Y^H}{w_H} \right) \geq 2v^2 + 2v^2 (L) - \xi \left( \frac{Y^L}{w_L} \right) - \xi \left( \frac{Y^H}{w_H} \right),$$

and

$$(Y^H - y^H) + (Y^L - y^L) + 2 (p - I) x^* + (q - I) (x^*(H) + x^*(L)) \geq 0$$

B.1 Second period goods

The first order condition with respect to $q'$ is

$$\left[ v^*_y (H) x^{*i} (H) + v^*_y (L) x^{*i} (H) \right] (1 - \mu) - 2 \mu v^*_y (L) x^{*2} (L) +$$

$$\lambda \left\{ - 2 (p - I) x^*_y s_i + x^* (H) + x^* (L) +$$

$$(q - I) \left[ x^*_y (H) + x^*_y (L) + (x^*_y (H) + x^*_y (L)) s_i \right] \right\} = 0 \tag{28}$$

where we used the fact that $v^*_y (k) = v^*_y (k) x^{*i} (k)$ for $k = H, L$.

Now, the first order conditions with respect to $y^H$

$$v^*_y (H) (1 + \mu) + \lambda \left[ - 1 - 2 (p - I) x^*_y s^*_H u +$$

$$(q - I) \left[ x^*_y (H) + x^*_y (L) \right] s^*_H u + x^*_y (H) \right] = 0, \tag{29}$$

and $y^L$

$$v^*_y (L) (1 + \mu) - \mu v^*_y (L) + \lambda \left[ - 1 - 2 (p - I) x^*_y s^*_L u +$$

$$(q - I) \left[ x^*_y (H) + x^*_y (L) \right] s^*_L u + x^*_y (L) \right] = 0. \tag{30}$$

Next, we multiply (29) by the vector $x^*_y (H)$ and (30) by $x^*_y (L)$ add the two, and replace in (28) to obtain

$$2 \mu v^*_y (L) \left[ x^{*i} (L) - x^{*2} (L) \right] + \lambda 2 (p - I) \left[ x^*_y \left[ s_i - \left( s_{y L} x^{*i} (H) + s_{y H} x^{*i} (L) \right) \right] +$$

$$\lambda (q - I) \left[ x^*_y (H) + x^*_y (L) \right] \left[ s_i - \left( s_{y H} x^{*i} (H) + s_{y L} x^{*i} (L) \right) \right] +$$

$$\left( \lambda 0 - l \right) \left[ \left( x^*_y (H) - x^*_y (H) x^{*i} (H) \right) + (x^*_y (L) - x^*_y (L) x^{*i} (L)) \right] = 0$$
B.2 First period goods

First order condition with respect to $p^d$ is

\[
2v^*_y x^d (1 + \mu) - 2v^2 x^d + \lambda \{2x^d + 2(p - 1) x^*_y - [2(p - 1)x^*_y - (q - 1)(x^*_y (H) + x^*_y (L))}\} s^*_j = 0
\]

using the facts that $v^*_y = -v^*_y x^d$ and $v^2 = -v^2 x^d$.

From the agent’s first order condition we know that $2v^*_y = v^*_y (H) + v^*_y (L)$ and $v^2 = v^2 (L)$. Therefore, we may write it as

\[
2\mu v^*_y [x^d - x^d] + \lambda (p - 1) x^*_y [s^*_j - (s_{yH} + s_{yL}) x^d] + \\
\lambda (q - 1) (x^*_y (H) + x^*_y (L)) [s^*_j - (s_{yH} + s_{yL}) x^d] + \\
2\lambda (p - 1) (x^*_j - x^*_y x^d) = 0
\]

Now it is the deviation in the consumption of $x^d$ in the first period that determines the discouragement of this good.

Using the analogous definitions, for $\delta^*_j$ and $\tilde{x}^*_j$, the above expression simplifies to

\[
\frac{\mu v^*_y}{\lambda} [x^d - x^d] = - \left[ (p - 1) x^*_y + 1(q - 1) \frac{x^*_y (H) + x^*_y (L)}{2} \right] \delta^*_j - (p - 1) \tilde{x}^*_j
\]
<table>
<thead>
<tr>
<th>Article Number</th>
<th>Title</th>
<th>Authors</th>
<th>Publication Date</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>458</td>
<td>DESEMPENHO DE ESTIMADORES DE VOLATILIDADE NA BOLSA DE VALORES DE SÃO PAULO</td>
<td>Bernardo de Sá Mota, Marcelo Fernandes</td>
<td>Outubro de 2002</td>
<td>37 págs.</td>
</tr>
<tr>
<td>461</td>
<td>DECENT WORK AND THE INFORMAL SECTOR IN BRAZIL</td>
<td>Marcelo Côrtes Neri</td>
<td>Novembro de 2002</td>
<td>115 págs.</td>
</tr>
<tr>
<td>462</td>
<td>POLÍTICA DE COTAS E INCLUSÃO TRABALHISTA DAS PESSOAS COM DEFICIÊNCIA</td>
<td>Marcelo Côrtes Neri, Alexandre Pinto de Carvalho, Hessia Guilhermo Costilla</td>
<td>Novembro de 2002</td>
<td>67 págs.</td>
</tr>
<tr>
<td>466</td>
<td>INFLAÇÃO E FLEXIBILIDADE SALARIAL</td>
<td>Marcelo Côrtes Neri, Maurício Pinheiro</td>
<td>Dezembro de 2002</td>
<td>16 págs.</td>
</tr>
<tr>
<td>467</td>
<td>DISTRIBUTIVE EFFECTTS OF BRAZILIAN STRUCTURAL REFORMS</td>
<td>Marcelo Côrtes Neri, José Márcio Camargo</td>
<td>Dezembro de 2002</td>
<td>38 págs.</td>
</tr>
<tr>
<td>468</td>
<td>O TEMPO DAS CRIANÇAS</td>
<td>Marcelo Côrtes Neri, Daniela Costa</td>
<td>Dezembro de 2002</td>
<td>14 págs.</td>
</tr>
<tr>
<td>469</td>
<td>EMPLOYMENT AND PRODUCTIVITY IN BRAZIL IN THE NINETIES</td>
<td>José Márcio Camargo, Marcelo Côrtes Neri, Mauricio Cortez Reis</td>
<td>Dezembro de 2002</td>
<td>32 págs.</td>
</tr>
</tbody>
</table>


477. INADA CONDITIONS IMPLY THAT PRODUCTION FUNCTION MUST BE ASYMPTOTICALLY COBB-DOUGLAS - Paulo Barelli; Samuel de Abreu Pessoa – Março de 2003 – 4 págs.


479. A NOTE ON COLE AND STOCKMAN - Paulo Barelli; Samuel de Abreu Pessoa – Abril de 2003 – 8 págs.


481. ON THE WELFARE COSTS OF BUSINESS CYCLES IN THE 20TH CENTURY - João Victor Issler; Afonso Arinos de Mello Franco; Osmani Teixeira de Carvalho Guillén – Maio de 2003 – 29 págs.


<table>
<thead>
<tr>
<th>ID</th>
<th>Title</th>
<th>Authors</th>
<th>Date</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>487</td>
<td>USING IRREGULARLY SPACED RETURNS TO ESTIMATE MULTI-FACTOR MODELS: APPLICATION TO BRAZILIAN EQUITY DATA</td>
<td>Álvaro Veiga; Leonardo Rocha Souza</td>
<td>Junho de 2003</td>
<td>26 pág.</td>
</tr>
<tr>
<td>489</td>
<td>CONVEX COMBINATIONS OF LONG MEMORY ESTIMATES FROM DIFFERENT SAMPLING RATES</td>
<td>Leonardo R. Souza; Jeremy Smith; Reinaldo C. Souza</td>
<td>Julho de 2003</td>
<td>20 pág.</td>
</tr>
<tr>
<td>490</td>
<td>IDADE, INCAPACIDADE E A INFLAÇÃO DO NÚMERO DE PESSOAS COM DEFICIÊNCIA</td>
<td>Marcelo Neri; Wagner Soares</td>
<td>Julho de 2003</td>
<td>54 pág.</td>
</tr>
<tr>
<td>492</td>
<td>THE MISSING LINK: USING THE NBER RECESSION INDICATOR TO CONSTRUCT COINCIDENT AND LEADING INDICES OF ECONOMIC ACTIVITY</td>
<td>João Victor Issler; Farshid Vahid</td>
<td>Agosto de 2003</td>
<td>26 pág.</td>
</tr>
<tr>
<td>493</td>
<td>REAL EXCHANGE RATE MISALIGNMENTS</td>
<td>Maria Cristina T. Terra; Frederico Estrella Carneiro Valladares</td>
<td>Agosto de 2003</td>
<td>26 pág.</td>
</tr>
<tr>
<td>494</td>
<td>ELASTICITY OF SUBSTITUTION BETWEEN CAPITAL AND LABOR: A PANEL DATA APPROACH</td>
<td>Samuel de Abreu Pessoa; Silvia Matos Pessoa; Rafael Rob</td>
<td>Agosto de 2003</td>
<td>30 pág.</td>
</tr>
<tr>
<td>495</td>
<td>A EXPERIÊNCIA DE CRESCIMENTO DAS ECONOMIAS DE MERCADO NOS ÚLTIMOS 40 ANOS</td>
<td>Samuel de Abreu Pessoa</td>
<td>Agosto de 2003</td>
<td>22 pág.</td>
</tr>
<tr>
<td>496</td>
<td>NORMALITY UNDER UNCERTAINTY</td>
<td>Carlos Eugênio E. da Costa</td>
<td>Setembro de 2003</td>
<td>08 pág.</td>
</tr>
<tr>
<td>498</td>
<td>REDISTRIBUTION WITH UNOBSERVED 'EX-ANTE' CHOICES</td>
<td>Carlos Eugênio E. da Costa</td>
<td>Setembro de 2003</td>
<td>30 pág.</td>
</tr>
<tr>
<td>499</td>
<td>OPTIMAL TAXATION WITH GRADUAL LEARNING OF TYPES</td>
<td>Carlos Eugênio E. da Costa</td>
<td>Setembro de 2003</td>
<td>26 pág.</td>
</tr>
<tr>
<td>500</td>
<td>AVALIANDO PESQUISADORES E DEPARTAMENTOS DE ECONOMIA NO BRASIL A PARTIR DE CITAÇÕES INTERNACIONAIS</td>
<td>João Victor Issler; Rachel Couto Ferreira</td>
<td>Setembro de 2003</td>
<td>29 pág.</td>
</tr>
<tr>
<td>501</td>
<td>A FAMILY OF AUTOREGRESSIVE CONDITIONAL DURATION MODELS</td>
<td>Marcelo Fernandes; Joachim Grammig</td>
<td>Setembro de 2003</td>
<td>37 pág.</td>
</tr>
<tr>
<td>502</td>
<td>NONPARAMETRIC SPECIFICATION TESTS FOR CONDITIONAL DURATION MODELS</td>
<td>Marcelo Fernandes; Joachim Grammig</td>
<td>Setembro de 2003</td>
<td>42 pág.</td>
</tr>
</tbody>
</table>