The Risk Properties of Human Capital and the Design of Government Policies.*

Carlos E. da Costa[†], *Getulio Vargas Foundation*
Lucas J. Maestri[‡], *Getulio Vargas Foundation*

September 17, 2004

Abstract

Whether human capital increases or decreases wage uncertainty is an open question from an empirical standpoint. Yet, most policy prescriptions regarding human capital formation are based on models that *impose* riskiness on this type of investment. In a two period and finite type optimal income taxation problem we derive prescriptions that are robust to the risk characteristics of human capital: savings should be discouraged, human capital investments encouraged and both types of investment driven to an efficient level from an aggregate perspective. These prescriptions are also robust to what choices are observed, even though the policy instruments used to implement them are not.

Keywords: Education; Optimal Taxation; Endogenous Policy Instruments
JEL Classification: H21, H24

---

*We thank Bev Dahlby, Luis Braido, Humberto Moreira and Daniel Gottlieb for their invaluable comments. All errors are our own.

†Corresponding author. 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: 55-21-2559.5493E-mail: cdacosta@fgv.br.

‡E-mail: lmaestri@fgvmail.br

1 Introduction

The close connection between risk characteristics of human capital investment and optimal tax and educational policy prescriptions dates back, at least, to the pioneering work of Eaton and Rosen (1980).

In a two period representation of life-cycle choices, Eaton and Rosen (1980) assume that human capital is risky, and that insurance markets for this type of life time uncertainty do not exist. This creates a role for the government to use tax instruments for insuring against this purely idiosyncratic risk. They derive prescriptions regarding labor income taxes and educational policy under the assumption that labor supply is inelastic. This assumption about agents' behavior, is an important caveat if one considers all possible margins—e.g., age of retirement, effort, etc.—that are available for adjustment after one has fully realized one's earnings potential. In a later contribution, Hamilton (1987) remedies this problem by including labor supply responses after realization of uncertainty. He shows that additional restrictions on preferences deliver results that are similar those found in Eaton and Rosen (1980).

Despite the fact that most tax prescriptions derived in these papers depend on restrictive assumptions on preferences, one cannot overemphasize the role they played in building our understanding of key issues of the problem. For our purposes, it is important to highlight some of the main conclusions, though, as we have mentioned, they should not be taken as general propositions, but rather as results derived for what may be regarded as reasonable cases. They are: *i)* proportional taxes are usually superior to lump-sum taxes with regards to its social welfare consequences; governments should *ii)* encourage human capital and *iii)* discourage savings. Underlying all these results is a single driving force: the risky nature of human capital investment.

When investing in human capital an agent affects not only her expected productivity, but also (potentially) the entire wage distribution. Roughly, if most of the increase in productivity takes place in the 'good' states of nature, i.e., when the marginal utility is lower, then human capital is risky. More importantly, although this type of risk is purely idiosyncratic, no private markets exist to insure against it. The consequence is under-investment in human capital formation from an aggregate perspective. The government may, then, increase the agents' expected utility by reducing human capital risk with a proportional tax and ameliorate the under-investment problem by directly subsidizing education.

The dependence of optimal policies on the risk characteristics of human capital has long been recognized in the literature. Hamilton (1987, p. 380), for example, writes: "The basic requirement for the result [social welfare is increasing in human capital at the optimum private choices] is that the marginal return to human capital remains high for favorable states of the world in which marginal utility of income is low." If, on the contrary, human capital returns are higher on the states of nature where marginal utility of income is higher, then the optimal policy prescriptions may be reversed, i.e., it may be optimal to reduce human capital investment from its private optimal level.

But can we take for granted the risky nature of human capital investment? Despite the existence of a large (and growing) literature exploring the issue, no conclusive evidence regarding which case is most relevant is yet available. The opposite assumption to the one used in the literature is, therefore, just as appealing.[1]

We shall not explore the empirical literature, referring instead to the discussion in the introduction of a recent article by Andersson and Anderberg (2003). We do so not only because we believe that we could not match the quality of their arguments but also, and more fundamentally, because the main point of our paper is exactly the fact that the risk characteristic of human capital is inessential to the qualitative properties of optimal policy.

We write a two periods model along the lines of Hamilton (1987) but allow for more powerful tax instruments: we consider a fully non-linear tax on labor income and possibly state dependent taxes on return of savings and human capital.

When allowing for this type of instruments in analyzing the optimal taxation problem, one must model human capital policy along the lines of Mirrlees' (1971) which involves dealing with a multi-period agency problem which was only starting to be explored when the first papers regarding human capital investments came out. It should come as no surprise that, even though agency approaches to deal with the problem are found as early as Ulph and Hare (1979), risk is altogether eliminated by their collapsing choices in a single period.[2] It was only recently that Naito (2004) has dealt with human capital formation within a dynamic agency framework, aiming at showing that differential taxation of goods remains optimal when there is comparative advantage in human capital formation.

Following Guesnerie (1995), we start with a description of the informational structure of the economy and derive tax instruments in two steps. First, we use the revelation principle to determine optimal allocations by means of a truthful revelation direct mechanism. Then, we characterize tax instruments capable of implementing such allocations.

[1] One clear example of the insurance properties of human capital is the fact that unemployment spells are found to be shorter for higher educated workers.

[2] The same procedure is followed by Brett and Weymark (2000).

As it turns out, few restrictions on preferences are needed in our framework for policy prescriptions to be derived. We are able to show that, if government has full control over private choices of savings and investment in human capital, savings should be set below and human capital investments above what would be privately chosen by agents *independently* of whether investment in human capital increases or reduces wage risk.

To understand what drives our result one must bear in mind the fact that, contrary to the earlier literature, we model the government optimization as a mechanism design problem. Risk sharing is taken care of by the non-linear income tax schedule. The role of other instruments is simply to relax the incentive compatibility constraints. It is not, then, the risk characteristics of human capital but rather how deviant (or the anticipation of deviant) behavior affects investment choices that ultimately determines optimal policies.

Another important characteristic of optimal policies is the prescription of aggregate efficiency for both types of investment: physical and human capital. On the one hand, agents private choices are distorted, while on the other, aggregate return is determined by a condition compatible with the prescription of no capital income taxation of Judd (1985), Chamley (1986) and Lucas (1990). This result is in contrast with the prescriptions found in Anderson and Anderberg (2003) where risk-enhancing (reducing) human capital should be set at a level which is lower (higher) than the first best one.

As for implementation, despite our having no heterogeneity at the time investment choices are made, simple linear taxes or subsidies are not capable of implementing the optimal allocations. One of the consequences is that education subsidies do not substitute for compulsory education as one might expect, given ex-ante homogeneity among agents. With regards to savings, the optimal policy is even harder to interpret since the government must induce lower savings. We then characterize implementation via taxes both for savings and human capital investments. Marginal tax rates are shown to depend on second period labor output, i.e., marginal taxes for both type of investments are stochastic and dependent on labor output.[3]

Implementation thus requires the use of instruments that are not currently observed and/or depend on apparently extreme assumptions on observability of choices. To overcome this caveat we relax the assumptions on the controllability of savings and human capital investment by the government, letting these choices be affected only by simple linear taxes or subsidies.

Encouragement of education and discouragement of savings still obtains albeit with

[3] This result on taxation of savings was first derived, to the best of our knowledge, by Albanesi and Sleet (2003) in a somewhat similar environment.

slightly different meaning. If the government policy is towards either savings or human capital formation but not both at the same time, then it is optimal to tax savings or subsidize education. When both instruments are used, the signs of tax rates are ambiguous, though the formulae resemble the encouragement/discouragement expressions derived in Mirrlees (1976): savings ought to be 'approximately discouraged' and human capital formation 'approximately encouraged', in a sense to be made precise in the main text.

The remainder of the paper is organized as follows. Section 2 described the model economy discussing the competitive equilibrium and the properties of the first best optimum—sections 2.1 and 2.2, respectively. Section 3 describes optimal taxation in a world where government has complete control of agents' investment choices, both in human capital and in a risk free technology. Implementation is discussed in section 3.1, where the tax systems that are needed to implement the allocations in section 3 are shown to be very far from what is currently observed. In section 4 we consider a more sensible assumption on observability of savings and human capital formation that means that the government may influence these choices but cannot exert direct control over them. In section 4.6 an intermediate case is explored. Section 5 concludes.

2 The Economy

The economy, based on Hamilton(1987), is populated by an atomistic measure of ex-ante identical agents with preferences represented by

$$u\left(c^{1}\right)+E\left[u\left(c^{2}\right)-\zeta\left(l\right)\right]$$

where, c^1 and c^2 are, respectively, first and second period consumptions and l is second period labor supply. For simplicity we assume that, in the first period, leisure is not valued. We assume u and ζ to be smooth functions such that $u' > 0, u'' < 0$, $\zeta' > 0$, $\zeta'' < 0$ $\lim_{c\searrow 0} u'(c) = \infty$ and $\lim_{l\nearrow 1} \zeta'(l) = \infty$.

Labor supply, or effort, produces efficiency units, Y, according to the relation $Y = wl$, where w is the agents productivity. As mentioned, we abstract from leisure choices in the first period and assume that agents are endowed with one efficiency unit that may be sold to firms or used in producing human capital h. Therefore labor supply in the first period is simply the difference between the time endowment, and time spent in human capital formation, $1-h$. As we shall see, this is also an agent's earnings in the first period, as well. We follow Eaton and Rosen (1980) and Hamilton (1987) in considering foregone earnings as the only costs of this type of investment. Including other costs would add notational burden and would not provide new insights to the problem.

Productivity is a function of human capital investment h and 'talent' θ. Uncertainty arises in this economy because, when still young, agents do not know their innate talent θ. That is, an agent who dedicates h units of time in the first period to human capital formation, will have $w(h,\theta^i)$ efficiency units of labor, which we shall call her productivity, in period two if her talent turns out to be θ^i. Here we take w to be a differentiable, strictly increasing (in the two arguments) and strictly concave function with $w(0,\theta) = 0$ and $\lim_{h \searrow 0} \partial w(h,\theta)/\partial h = \infty$ for all θ.

Our model follows Mirrlees (1971) in considering that, once uncertainty is resolved, talent is privately information, although the distribution function of talent is common knowledge.

Firms transform efficiency units into consumption goods using a linear technology with units normalized in such a way that one efficiency unit produces one unit of consumption. Firms cannot observe l or w, therefore, labor contracts are defined in terms of efficiency units $Y \equiv wl$ which are observed.

The economy is competitive in the usual sense. Individuals maximize utility and firms maximize profits taking prices—wages and interest rates—as given. The zero profit condition for firms implies that one efficiency unit is paid its marginal product of 1. Also important is the fact that firms compete for workers in every period, and that agents cannot commit to working for the same firm for the two periods.[4] This assumption precludes insurance contracts between firms and workers and introduces a role for optimal taxation.

Before writing the agents problem, we add to the model the fact that in period 1 agents need not consume their income, $1-h$. Borrowing and lending, denoted with s (positive if lending and negative if borrowing), is possible at a gross interest rate of 1. This assumption of capital market efficiency is not made for sake of realism, but to focus on the role of risk properties of human capital in policy design.[5] As we shall see, this type of market failure should aggravate the problem and make our case even stronger.

Hence, absent taxes, the budget constraints for the agents are

$$c^1 \leq 1 - h - s,$$

in the first period, and

$$c^i \leq l^i w(h,\theta^i) + s,$$

[4] Our model is a representation of life-cycle choices for which this type of commitment doesn't seem to be a good representation of actual labor contracts. Elsewhere—e.g., Golosov and Tsyvinsky (2003b)—a competitive equilibrium where these long term contracts are possible is considered.

[5] A very interesting discussion of the role of optimal taxation in a market where the absence of collateral impedes the financing of education is found in Hoff and Lyon (1995).

in the second period, state of nature i. Here the superscript i in l^i denotes the fact that labor supply is state contingent—it is only decided after the realization of uncertainty.

One should not, however, view second period labor supply responses as simple variation in hours which, as pointed out by Eaton and Rosen (1980), seems not to be very responsive to changes in real after tax wages. Rather, this accounts for adjustments in long term labor/leisure choices after an individual has fully realized her earnings capacity.

In what follows, we shall assume that there are only two states of the world, H and L with $\theta^H > \theta^L$. We also adopt the 'law of large numbers convention' that the θ's are independently and identically distributed, and that the number of individuals composing the population is sufficiently large to eliminate aggregate uncertainty. Finally, and just for notational convenience, we assume that types are in equal proportion in this economy.

2.1 The No-intervention Equilibrium

In this section, we consider a no-intervention regime, leaving for the next section the case of a benevolent government which will use the instruments available to increase expected utility.

Following the description of the economy in section 2, the timing of the model is as follows. In the first period, every worker decides how to split her time between work and investment in human capital formation. She is paid the total efficiency units she supplies, $1-h$, and chooses the first period consumption and, consequently, her savings, s which may be negative.

In the second period, conditional on the variables chosen in the first period and on her type (now revealed), she has to choose how many efficiency units, $Y(h, s, \theta)$, to supply.

As is usual we solve agents optimization problem backwards. That is, we start by defining the second period indirect utility functions,

$$V\left(\theta^i, s, h\right) \equiv \max_{Y^i} u\left(Y^i + s\right) - \zeta\left(\frac{Y^i}{w(h, \theta^i)}\right), \quad i = H, L \tag{1}$$

and then solving the first period optimal choices,

$$\max_{s,h} u\left(1-h-s\right) + \frac{1}{2}\left[V\left(\theta^H, s, h\right) + V\left(\theta^L, s, h\right)\right].$$

The first order conditions, which under our assumptions are necessary and sufficient for a maximum, yield,

$$2u'\left(c^1\right) = u'\left(c^H\right) + u'\left(c^L\right)$$

and

$$2u'\left(c^1\right) = \zeta'\left(l^H\right) l^H \frac{w_h(h, \theta^H)}{w(h, \theta^H)} + \zeta'\left(l^L\right) l^L \frac{w_h(h, \theta^L)}{w(h, \theta^L)}.$$

where $w_h(h,\theta^i) \equiv \partial w(h,\theta^i)/\partial h$.

To further explore this expression, we use the first order conditions for problem (1) to write.

$$2u'\left(c^1\right) = u'(c^H)l^H w_h(h,\theta^H) + u'(c^L)l^L w_h(h,\theta^L)$$

dividing both sides by $u'\left(c^H\right) + u'\left(c^L\right)$ we have,

$$1 = \pi^H w_h(h,\theta^H)l^H + \pi^L w_h(h,\theta^L)l^L \lesseqgtr \frac{1}{2}[w_h(h,\theta^H)l^H + w_h(h,\theta^L)l^L]$$

where π^i is the risk-adjusted probability of state i. On the right hand side of the expression above we have the marginal benefit of the investment in human capital formation in terms of efficient units. For sufficiently low labor supply elasticities there is under-investment (over-investment) if $w_{h\theta}(h,\theta) > 0$ ($w_{h\theta}(h,\theta) < 0$), where $w_{h\theta}(h,\theta) \equiv \partial^2 w(h,\theta^H)/\partial h \partial\theta$.

In Eaton and Rosen (1980) and Hamilton (1987) education yields higher returns in the states of nature where people are more productive.[6] That is, private returns are higher when the marginal utility of income is lower. Because there are no insurance markets for this risk, there is under-investment in human capital. As a consequence, introducing a small subsidy on education improves welfare by reducing under-investment.

It is then clear, that these results crucially depend on the fact that education is risky, which in turn depends on the sign of the cross derivatives and elasticity of labor supply. Andersson and Anderberg (2003) have formalized the idea—already recognized in Eaton and Rosen (1980) and Hamilton (1987)—that different assumptions on technology and labor supply would lead to different policies. If education decreases wage variance, an equally plausible assumption, and there would be over-investment, thus reversing tax prescriptions.

This is very discomforting. The qualitative properties of optimal policies cannot be assessed. Because there is no strong evidence regarding which assumption regarding the risk characteristics of human capital investment best adheres to the data, one cannot characterize optimal policies. As we shall see, however, our results do not depend on any of these issues. Prescriptions are robust to changes in risk characteristics of human capital.

First we describe the first best solution which will provide some other important grounds for comparing the prescriptions we shall derive.

2.2 First Best Allocations

The purpose of this section is to provide a benchmark against which the results in other sections may be compared.

[6] They use a specification that guarantees $\partial^2 w/\partial h\partial\theta > 0$ and assume that labor supply is inelastic or impose conditions that guarantee that labor supply will respond in the 'right' direction.

We characterize optimal savings and investment in human capital by solving the program

$$\max_{s,h,c^1,(c^i,l^i)_{i=H,L}} 2u\left(c^1\right) + \sum\nolimits_{i=H,L} \left[u\left(c^i\right) - \zeta\left(l^i\right)\right]$$

$$\text{s.t.} \qquad \sum\nolimits_{i=H,L} \left[c^i - l^i w\left(h, \theta^i\right)\right] \leq 2\left(1 - h - c^1\right) \qquad [\,\lambda\,]$$

It is easy to see that combining the first order conditions of this problem one gets labor supply efficiency,

$$u'\left(c^i\right) w^i(h, \theta^i) = \zeta'\left(l^i\right), \tag{2}$$

consumption smoothing across time and states of nature,

$$u'\left(c^1\right) = u'\left(c^H\right) = u'\left(c^L\right), \tag{3}$$

and efficiency in human capital investment,

$$1 = \frac{1}{2}\left[l^H w_h(h, \theta^H) + l^L w_h(h, \theta^L)\right]. \tag{4}$$

The last expression, i.e., the optimality condition for human capital investment, illustrates the fact that the marginal gain from an extra unit of time spent in enhancing human capital—as measured by the right hand side of (4)—is simply the expected increase in the output at the social optimum level of labor supply.

This condition is important because, throughout the paper, over and under-investment is always a statement about how the particular value compares to the one determined in (4).

Notice also that a first best allocation requires equating marginal utility of income in both states of nature, as shown in (3). It also means that more productive agents work more—as shown in (2)—thus having a lower utility than less productive agents.

This naturally means that if information about productivity is private—as is assumed throughout the paper—the first best allocation is not be implementable. Agents who turn out to be more productive just announce that they are less productive and get more utility then what they get if they announce truthfully.

3 Optimal Taxation

In this section, we assume that the government can fully control all the investment choices made by the agents, that including investments in human capital and savings decisions. Later we discuss what this assumption implies for tax instruments and the consequences of restricting the set of instruments available to the government.

Though the government can directly choose h and s, the first best allocation is still unfeasible, since only efficiency units—the traded objects—and not hours supplied are observed.

We write the mechanism design problem faced by the government and interpret it as an optimal tax system.

$$\max_{h,c^1,(c^i,Y^i)_{i=H,L}} 2u\left(c^1\right)+\sum\nolimits_{i=H,L}\left[u\left(c^i\right)-\zeta\left(\frac{Y^i}{w(h,\theta^i)}\right)\right]$$

subject to

$$u\left(c^H\right)-\zeta\left(\frac{Y^H}{w(h,\theta^H)}\right)\geq u\left(c^L\right)-\zeta\left(\frac{Y^L}{w(h,\theta^H)}\right) \qquad [\ \mu\ ]$$

and

$$\sum\nolimits_{i=H,L}\left(c^i-Y^i\right)\leq 2\left(1-h-c^1\right) \qquad [\ \lambda\ ]$$

We shall concentrate on two issues in particular: the second best choices of savings and investment in human capital.

The first order conditions with respect to c^1, c^H and c^L are, respectively, $u'\left(c^1\right)=\lambda$, $u'\left(c^H\right)(1+\mu)=\lambda$, and $u'\left(c^L\right)(1-\mu)=\lambda$. Therefore,

$$2u'\left(c^1\right)=\left[u'\left(c^H\right)+u'\left(c^L\right)\right]+\mu\left[u'\left(c^H\right)-u'\left(c^L\right)\right]. \qquad (5)$$

The right hand side of (5) is less then $u'\left(c^H\right)+u'\left(c^L\right)$ if and only if $c^H>c^L$. But this is exactly the case as one can show by usual single-crossing arguments. This result is reminiscent of Cremer and Gahvari (1995), da Costa and Werning (2000) and Golosov et al. (2003a).

The next result is, as far as we know, new in such a context.

$$\sum_{i=H,L}\zeta'\left(l^i\right)l^i\frac{w_h(h,\theta^i)}{w(h,\theta^i)}-2u'(c^1)=-\mu\frac{w_h(h,\theta^H)}{w(h,\theta^H)}[\zeta'\left(l^H\right)l^H-\zeta'\left(l^{L|H}\right)l^{L|H}] \qquad (6)$$

where l^i $(i=H,L)$ is the labor supply of agent type i and $l^{L|H}\equiv Y^L/w(h,\theta^H)$ is the labor supply of a high type agents who announces to be of a low type.

The first term in the left hand side is simply the private marginal benefit of education given by the decrease in marginal expected effort conditional on a given level of required efficiency units per type. The second term is the marginal cost. Absent any government policy this would be equal to 0.

The right hand side is, in general, different from 0. In fact, because $l^{L|H}<l^L$ and ζ is strictly convex, (6) implies

$$\frac{w_h(h,\theta^H)}{w(h,\theta^H)}\zeta'\left(l^H\right)l^H+\frac{w_h(h,\theta^L)}{w(h,\theta^L)}\zeta'\left(l^L\right)l^L<2u'\left(c^1\right). \qquad (7)$$

The government creates a wedge between the private marginal costs and benefits of investment in human capital proportional to the difference between the marginal benefit of this form of investment for high productivity agents who announce truthfully and those who lie. Choices are thus distorted in the direction of hurting off-equilibrium behavior. This result *does not* depend on the sign of $w_{h\theta}(h, \theta^H)$!

The point we stress here is that, in deriving optimal policies we do not compare choices made in autarchy with first best choices, but choices of those who intend to deviate from truthful announcement with choices of those who intend to comply. Because investments in human capital is lower for the former than for the latter, the optimal policy requires encouraging human capital formation.

To further characterize optimal policies, we need the other two first order conditions,

$$\zeta'\left(l^H\right)l^H\left(1+\mu\right)=\lambda Y^H, \tag{8}$$

and

$$\zeta'\left(l^L\right)l^L-\mu\zeta'(l^{L|H})l^{L|H}=\lambda Y^L. \tag{9}$$

From (6), (8) and (9) we have

$$u'\left(c^1\right)=\frac{\lambda}{2}\left[w_h(h,\theta^H)l^H+l^Lw_h(h,\theta^L)\right]\Rightarrow 1=\frac{1}{2}\left[w_h(h,\theta^H)l^H+l^Lw_h(h,\theta^L)\right]$$

which is exactly (4). This is in contrast with the competitive equilibrium where uncertainty concerning the returns of education leads to a level of investment in human capital formation that is generally non-optimal from an aggregate perspective.

Here, however, the optimal level of human capital investment is reached at the point where aggregate marginal costs and benefits of education are equal. This is in contrast with Anderson and Anderberg (2003). They find that if human capital increases wage risk, then the second best investment in human capital should be set at a level that is lower than the first best one.[7] The opposite holds when human capital reduces wage risk.

3.1 Implementation

As we have seen, the government distorts choices of savings and human capital investment, inducing—from a private perspective—less savings and more investment in human capital formation. How does this wedge translate into government policies? This is the question we tackle along this section.

[7] The statement needs some qualification since the optimal level of h need not coincide with the first best one since labor supply differs, in general.

3.1.1 Compulsory Choices and Taxation

If government can force agents to choose specific levels of human capital and savings then the answer to our problem is trivial. Schooling would be compulsory at the optimal level. In practice, compulsory education usually means compelling that a certain amount of time is spent in 'school', but other dimensions of education such as effort may not be enforceable.

For savings, the story is a little different because the government must induce under-savings. This idea is at odds with most policies adopted in modern economies: it implies that government should forbid people to save!

How can tax policies induce the optimal behavior? Since there is no heterogeneity at the moment where agents make their investment choices, a natural candidate policy is a linear tax on investment returns. In fact, a single tax is able to guarantee that, for every agent, $u'\left(c^1\right) = E\left[u'\left(c^*\right)\right](1-\tau)$, where c^1 and c^* are the equilibrium consumptions in first and second periods, respectively.

It turns out, however, that this is not enough. One has to be concerned with what is known as a *double deviation*—see Albanesi and Sleet (2003) and Kocherlakota (2004). Agents, anticipating the fact that they will lie about their productivities in the second period, choose a level of savings compatible with these off-equilibrium choices.

The tax system must, then, guarantee that the equilibrium savings choice is optimal not only along the equilibrium path but also off-equilibrium. The structure of such tax functions is discussed in 3.1.2. As we shall make clear, these will be dependent on second period choices and will differ across agents. Therefore, the observation of individual savings choices is crucial. In section 4, we relax this observability assumption and show the consequences for optimal policies.

3.1.2 Implementation via Taxes

We consider a tax schedule that is (possibly) dependent not only on the level of savings but also on investment in human capital and on an agents announcement. We discuss in some detail whether one can dispense with any of these features.[8]

The first thing we note is that, because one is allowing for fully non-linear taxes one cannot map wedges into taxes. For savings to be identical for liars and abiders—and this is what 'controlling savings' is all about—the marginal tax rates must be such that the

[8] The discussion on savings draws heavily on Albanesi and Sleet (2003), Golosov and Tsyvinsky (2003), Kocherlakota (2003) and Kocherlakota (2004).

marginal benefit of human capital investments for an agent along the equilibrium path,

$$u'\left(c^{H}\right)\left(1-\tau_{s}\left(s,h,H\right)\right)+u'\left(c^{L}\right)\left(1-\tau_{s}\left(s,h,L\right)\right) \tag{10}$$

must be identical to the marginal benefit for the same investments off the equilibrium path,

$$2u'\left(c^{L}\right)\left(1-\tau_{s}\left(s,h,L\right)\right). \tag{11}$$

Because $u'\left(c^{H}\right)<u'\left(c^{L}\right)$, (10) and (11) can only hold simultaneously if $\tau_{s}\left(s,h,H\right)<\tau_{s}\left(s,h,L\right)$—marginal taxes on savings must be state dependent, i.e., tax schedules must be non-separable between savings and announcements. Moreover, (10) and (11) imply $E\left[u'\left(c\right)\tau_{s}\right]>0$. But, because u' and τ_{s} are co-monotonic, $E\left[u'\left(c\right)\tau_{s}\right]>0$ need not imply $E\left[\tau_{s}\right]>0$.

The same is true with respect to human capital formation, though interpretation is a little harder here.

Implicit in our formulation is a tax schedule of the following kind. An agent chooses in period one a certain level of human capital investment, h. Then, in the second period, conditional on her announcement, and her savings, she pays a tax $\tau\left(s,h,Y\right)$. Her first order condition for human capital investment is

$$2u'\left(c\right)-\sum_{i=H,L}\zeta'\left(l^{i}\right)l^{i}\frac{w_{h}(h,\theta^{i})}{w(h,\theta^{i})}=-\left[\sum_{i=H,L}\tau_{h}\left(s,h,i\right)u'\left(c^{i}\right)\right], \tag{12}$$

if she plans to announce truthfully, and

$$2u'\left(c\right)-\frac{w_{h}(h,\theta^{L})}{w(h,\theta^{L})}\zeta'\left(l^{L}\right)l^{L}-\frac{w_{h}(h,\theta^{H})}{w(h,\theta^{H})}\zeta'\left(l^{L|H}\right)l^{L|H}=-2\tau_{h}\left(s,h,L\right)u'\left(c^{L}\right), \tag{13}$$

if she intends to lie.

From (12) and (13) it is immediate that

$$\frac{w_{h}(h,\theta^{H})}{w(h,\theta^{H})}\left[\zeta'\left(l^{H}\right)l^{H}-\zeta'\left(l^{L|H}\right)l^{L|H}\right]=\tau_{h}\left(s,h,H\right)u'\left(c^{H}\right)-\tau_{h}\left(s,h,L\right)u'\left(c^{L}\right)$$

The left hand side is positive. It must also be the case that

$$\tau_{h}\left(s,h,H\right)u'\left(c^{H}\right)+\tau_{h}\left(s,h,L\right)u'\left(c^{L}\right)<0 \tag{14}$$

for (7) to be satisfied. Which means that $\tau_{h}\left(s,h,L\right)<0$.

In sections 2.1 and 2.2 we have shown how it is usually argued that because human capital is a risky asset agents under-invest on it. Paradoxically, taxes that induce more risk by setting lower marginal tax rates for more productive agents cannot be ruled out here.

Although we are not able to sign expected marginal tax rates on the so-called empirical measure, we are able to sign them on the risk-neutral measure. Direct observation of 14 and 5 shows that $E^Q(\tau_s) > 0$ and $E^Q(\tau_h) < 0$. Optimal wedges are created by having expected risk-adjusted marginal tax rates to have the expected sign.

The fact that the results obtained thus far depend on tax instruments that do not resemble most current policies may be regarded in two different ways. First, one may just say that instruments currently in use are not optimal. Alternatively, our assumptions may not be very realistic. In particular, we have been assuming that the government can observe education and savings. First, one may argue that observable measures of human capital formation as years of schooling may be very far from fully capturing what is meant by education. Second, savings choices are at least partly non-observable by the government. Hence, in practice, the government may be restricted to using simple instruments like linear subsidies or taxes for both types of investment.

This type of consideration could cast some doubt about the importance of the results presented so far. Instead of dwelling upon this issue, we consider an environment where these choices are not observed, and show in the next section that results herein are robust—in a sense to be made precise—to restrictions in tax instruments associated to this (possibly) more realistic environment.

4 The Role of Unobserved Choices

If neither savings nor human capital investment is observed by the government, one can no longer define the agency problem as in the previous section, having to rely, instead, in a procedure that allows for these unobserved choices.

As usual, we start by describing the second period and working backwards. Let s^* and h^* denote, respectively, the equilibrium choices of savings and human capital which we will explain in detail shortly after this. Then, an allocation that is implemented by a truthful direct mechanism must satisfy

$$u(s^* + y^i) - \zeta\left(\frac{Y^i}{w(h^*, \theta^i)}\right) \geq u(s^* + y^j) - \zeta\left(\frac{Y^j}{w(h^*, \theta^i)}\right) \quad i, j = H, L. \tag{15}$$

This guarantees that once uncertainty is revealed, no matter what type she turns out to be, the agent finds in her best interest to reveal her true type.

Though necessary, conditions (15) do not suffice, in this case. Of concern here is again the *double deviation* problem. Even if, at the equilibrium levels of savings and human capital

s^* and h^*, the agent prefers the allocation intended for her, there may be alternative choices of h and s at which (15) does not hold and that are ex-ante preferred by the agent.

To deal with this, we follow da Costa (2004) by considering that, in the first period, agents conceive contingent plans, or strategies, that prescribe for each realization of θ what they ought to announce as being their innate ability. A strategy is, in this sense, a mapping $\sigma : \{H, L\} \to \{H, L\}$ from types to announcements.

This is not all that an agent does in the first period, however. She also chooses how much to save, s, and how much to invest in human capital, h. This may result in the problem being very complicated, in general. It turns out, however, that in our setup, the optimization problem related to savings and human capital investment is strictly convex *conditional on a given strategy*. Hence, associated to each strategy, σ^i, is a pair (h^i, s^i) of optimal ex ante choices. Because we only have two types, there are only four possible strategies, and consequently triples, $\left(\sigma^i, h^i, s^i\right)$, to compare.

The revelation principle guarantees that we can restrict our attention to allocations that can be implemented by a direct mechanism. Therefore, in characterizing the optimal tax schedule we choose allocations that induce the truthful announcement strategy, which we shall denote with a star, $\sigma^*(j) = j$, $j = H, L$.

Let σ^o represent an arbitrary alternative strategy. An allocation is implementable if (15) holds and if

$$2u\left(1 - s^* - h^*\right) + \sum_i \left\{ u(s^* + y^i) - \zeta\left(\frac{Y^i}{w(h^*, \theta^i)}\right)\right\} \geq$$
$$2u\left(1 - s^o - h^o\right) + \sum_i \left\{ u(s^o + y^{\sigma^o(i)}) - \zeta\left(\frac{Y^{\sigma^o(i)}}{w(h^o, \theta^i)}\right)\right\}, \tag{16}$$

for all σ^o, where s^o and h^o are, respectively, the optimal savings and investment corresponding to strategy σ^o.

In words, the allocation must be such that the truthful strategy is preferable to the other strategies allowing for the optimization with respect to h and s, which is equivalent to saying that the triple (σ^*, s^*, h^*) is better than any of the other three triples $(\sigma^o(\cdot), s^o, h^o)$.

There are five different incentive compatibility—henceforth IC—constraints: three in the first period (16) and two in the second (15). What we shall do next is to show that only one of the five IC constraints binds at the optimum.

4.1 Some Auxiliary Results

In the next few paragraphs we collect a series of results that will be necessary for defining the government's program.

The first proposition guarantees that once the first period IC constraints (15) are imposed one need not be concerned about second period IC constraints.

Proposition 1 *If constraints (16) are satisfied, so are constraints (15).*

The intuition for this result is rather simple. Assume that one of the second period IC constraints is binding. In this case, *at the equilibrium level of savings and human capital investment* the agent is indifferent between truthful announcement and adopting the strategy corresponding to this binding IC constraint. Re-optimizing s and h will break the indifference in favor of the alternative strategy thus violating one of the first period IC constraints. As a consequence, if (16) are satisfied, so are (15).

Next, we present a lemma that will be needed in most results that follow. It says that implementable allocations are increasing in types, i.e., then usual monotonicity requirement extends to our framework.

Lemma 1 *For an allocation* $\left(Y^i, y^i\right)_{i=H,L}$ *to be implementable it is necessary that* $\left(Y^H, y^H\right) \geq \left(Y^L, y^L\right)$. *If* $\left(Y^H, y^H\right) \neq \left(Y^L, y^L\right)$ *then* $\left(Y^H, y^H\right) >> \left(Y^L, y^L\right)$.

The next result guarantees that the strategy of always announcing falsely—$\sigma^o(j) = i, \;\; i \neq j$—cannot be optimal. More precisely, if there is a choice of (s, h) that makes this strategy better than truthful announcement, then there are also levels of savings and investment in human capital that make one of the two other alternative strategies optimal. Therefore, imposing the two IC constraints associated to strategies $\sigma^o(j) = L \;(j = H, L)$ and $\sigma^o(j) = H \;(j = H, L)$ guarantees that the third is also satisfied.

Proposition 2 *If the truthful strategy is no worse than strategies* $\sigma^o(j) = L \;(j = H, L)$ *and* $\sigma^o(j) = H \;(j = H, L)$*, then it is no worse than strategy* $\sigma^o(H) = L$*,* $\sigma^o(L) = H$*.*

We are down to two IC constraints only. We shall now show that the strategy of always announcing to be of a high productivity type—$\sigma^o(j) = H$, $j = H, L$—cannot be optimal, if the tax system is welfare increasing. This will bring us down to only one relevant IC constraint.

Notice that we rule out this strategy on different grounds than the other strategy. There are implementable allocations for which the relevant IC constraint is associated to the strategy of always announcing H.

The point here is that, though there might be allocations that are implementable at which the agents are indifferent between the truthful strategy and the strategy of always announcing high, these allocations yield a lower expected utility than what an agent gets in autarchy. Hence it should not bind at the optimum, given the objective of the government.

Proposition 3 *If the agent is indifferent between always telling the truth and always announcing H then the implementable allocation must result in a lower expected utility than what the agents gets in the competitive equilibrium.*

To summarize this section, we have shown that, under our assumptions, one needs only be concerned with one alternative strategy, $\sigma^o(j) = L$; $j = H, L$, which amounts to the agent always announcing to be of type L.

We are now in a position to write down the government's program.

4.2 The Optimal Taxation Problem

The government maximizes the agents' expected utility subject to the economy's resource constraint

$$\sum_{i=H,L} \left(Y^i - y^i\right) \geq 0$$

and, as shown in section 4.1, a single IC constraint

$$2u\left(1 - s^* - h^*\right) + \sum_{i=H,L} \left[u(s^* + y^i) - \zeta\left(\frac{Y^i}{w(h^*, \theta^i)}\right)\right] \geq$$
$$2\left[u\left(1 - s^o - h^o\right) + u(s^o + y^L)\right] - \sum_{i=H,L} \zeta\left(\frac{Y^L}{w(h^o, w^i)}\right).$$

To find optimal taxes we write the Lagrangian for the government's program,

$$\begin{aligned}\mathcal{L} \equiv\ & (1+\mu)\left\{2u\left(1 - s^* - h^*\right) + \sum_{i=H,L} \left[u(s^* + y^i) - \zeta\left(\frac{Y^i}{w(h^*, \theta^i)}\right)\right]\right\} - \\ & \mu\left\{2\left[u\left(1 - s^o - h^o\right) + u(s^o + y^L)\right] - \sum_{i=H,L} \zeta\left(\frac{Y^L}{w(h^o, \theta^i)}\right)\right\} + \lambda \sum_{i=H,L} \left(Y^i - y^i\right)\end{aligned}$$

From the first order conditions it is immediate that the marginal tax rate of a high productivity agent is found by exploring

$$\frac{\zeta'\left(l^H\right)}{u'\left(s^* + y^H\right)} = w(h^*, \theta^H).$$

Because the marginal rate of substitution between labor supply and consumption is exactly the price of leisure, the most productive agent is not distorted at the margin. The zero marginal tax rate result is preserved, here.

As for the low productivity agent, we have

$$\frac{\zeta'\left(l^L\right)}{u'\left(s^* + y^L\right)} \frac{1}{w(h^*, \theta^L)} = \left\{\lambda + \mu \left[\frac{\zeta'\left(Y^L / w(h^o, \theta^L)\right)}{w(h^o, \theta^L)} + \frac{\zeta'\left(Y^L / w(h^o, \theta^H)\right)}{w(h^o, \theta^H)}\right]\right\} \left\{\lambda + 2\mu u'\left(s^o + y^L\right)\right\}^{-1}$$

The left hand side of this expression is the marginal rate of substitution between y^L and Y^L, which is equal to one and there is no distortion. Therefore, we need to evaluate whether the right hand side of this expression to determine the sign of the marginal tax rate faced by the low productivity agent.

We first notice that the expression in the right hand side can be rewritten as

$$\left\{\frac{\phi}{\phi+1} + \frac{\mu}{2\left(\phi+1\right)} \left[\frac{\zeta'\left(Y^L / w(h^o, \theta^L)\right)}{u'\left(s^o + y^L\right) w(h^o, \theta^L)} + \frac{\zeta'\left(Y^L / w(h^o, \theta^H)\right)}{u'\left(s^o + y^L\right) w(h^o, \theta^H)}\right]\right\},$$

where $\phi \equiv \lambda / 2u'\left(s^o + y^L\right)$.

The term in brackets is the average marginal rate of substitution between leisure and consumption for an agent off the equilibrium path. Hence, the marginal tax rate faced by a low productivity agent will be positive or negative depending on whether the average marginal tax rate faced by an agent off the equilibrium path is positive or negative.

Though the rule differs from the case where no first period unobservable choices are present, the principle that drives tax prescriptions is the same. The idea is that if one is at the maximum redistribution that preserves Pareto efficiency, distorting choices may be useful in relaxing IC constraints if agents have different marginal rates of substitution between y and Y at this point.

When all choices are controlled by the government, we need only to compare the low type with the high type claiming to be a low type. Single-crossing is used to sign the marginal tax rate. Here, what must be considered is the average marginal rate of substitution of an agent who opts for the strategy of always announcing to be of type low, with that of a low type at the equilibrium choices of s and h. Unfortunately, we cannot resort to single-crossing to sign the marginal tax rate.

To further understand the issues at stake, we need to explore off equilibrium choices, starting from the next proposition.

Proposition 4 *At the optimum, $s^* < s^o$ and $h^* > h^o$.*

The intuition for this behavior is rather simple. When deciding to deviate by always announcing to be of a low productivity type, an agent anticipates less income and more leisure. Under our assumptions on preferences, this implies higher marginal utility of consumption and lower disutility of labor.

A higher marginal utility of income and a lower marginal disutility of labor makes it optimal to increase consumption and diminish leisure by choosing to save more and to get less education.

4.3 Educational Policy

In this section we assume that the government can introduce an anonymous tax (or subsidy) on human capital formation. To model it, we write the cost of education in terms of foregone earnings as $(1+\tau)\,h$ where τ is the tax on human capital formation.

We differentiate the Lagrangian with respect to τ at $\tau = 0$ to get

$$\begin{aligned} \left.\frac{\partial \mathcal{L}}{\partial \tau}\right|_{\tau=0} &= -(1+\mu)\,u'\left(c_1^*\right)h^* + \mu u'\left(c^1\right)h^o + \lambda h^* \\ &= -(1+\mu)\,h^*\left[u'\left(s^*+y^H\right)+u'\left(s^*+y^L\right)\right] + 2\mu u'\left(s^o+y^L\right)h^o + \lambda h^* \\ &= 2\mu u'\left(s^o+y^L\right)\left(h^o-h^*\right) < 0. \end{aligned} \tag{17}$$

where in the last step the first order conditions with respect to y^H and y^L were used.

This shows that, around $\tau = 0$, it is optimal to reduce taxes on (thus subsidizing) human capital formation. Eaton and Rosen (1980) and Hamilton (1987) derive an analogous result for taxation on wages based on assumptions which are a little more restrictive. In particular, it is worth pointing out that their result has a different interpretation. They show that a small tax rate on wages (equivalently a small subsidy for education) is welfare increasing because it provides insurance for the uncertain return of education. We, on the other hand, show that in an optimized economy this follows because taxing wages, or financing education, one can relax the incentive-compatibility constraint.

This result has a straightforward interpretation. From proposition 4, one can see that those who choose to cheat also decide to work more in the first period and get less education, since this strategy contemplates always announcing to be lower type. Hence, by subsidizing education the government can hurt more those who in the second period decide to lie about their types.

4.3.1 Optimal Taxes

At $\tau \neq 0$ things are not so straightforward. The problem here is that one can no longer guarantee which IC constraint binds at the optimum for arbitrary values of τ. Yet, we can still show that, *if* the binding IC constraint is the natural one, then it is optimal to subsidize education.[9]

When taxes are non-zero, government's budget constraint becomes

$$\sum_{i=H,L} \left(Y^i - y^i\right) + 2\tau h \geq 0$$

and equation (17) becomes

$$\frac{\partial \mathcal{L}}{\partial \tau} = 2\mu u' \left(s^o + y^L\right)(h^o - h^*) + 2\tau\lambda \frac{\partial \hat{h}^*}{\partial \tau},$$

where

$$\frac{\partial \hat{h}^*}{\partial \tau} \equiv \frac{\partial h^*}{\partial \tau} - \frac{1}{2}\left(\frac{\partial h^*}{\partial y^H} + \frac{\partial h^*}{\partial y^L}\right) h^* < 0$$

is a form of compensated investment in human capital.[10]

Under these conditions

$$\frac{\partial \mathcal{L}}{\partial \tau} = 0 \Rightarrow \tau < 0,$$

which means that optimal tax on education is negative: the policy prescription is to subsidize education.

In our model, there is no credit rationing nor externalities in the educational investment. Yet, we conclude that the optimal tax rate on education should be negative. This is not directly due to the fact that there is wage uncertainty in our model. Instead, wage uncertainty makes necessary an optimal taxation mechanism to redistribute wealth.

4.4 Taxing Investments

Defining t as the tax on investments, an analogous procedure to the one used in section 4.3 yields

$$\begin{aligned} \left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=0} &= -(1+\mu) \sum_{i=H,L} \left[u'(s^* + y^i)s^* - \mu 2u'(s^o + y^L)s^o\right] + \lambda s^* \\ &= \mu 2u'(s^o + y^L)\left[s^o - s^*\right] > 0. \end{aligned}$$

[9] Da Costa (2004) offers some arguments as of why we should expect this to be the case for any reasonable parametrization of the model.

[10] This extra term results from the fact that the first order conditions with respect to y^i $(i = H, L)$ have now an extra term, $\lambda\tau\partial h/\partial y^i$, that must be considered when proceeding with the same substitutions that were used in the previous section.

Once again we have the result that, around $t = 0$, it is optimal to increase taxes on (thus taxing) returns on investments.

4.4.1 Optimal Taxes

This case is exactly analogous to the case studied in section 4.3.1. Once again, assuming that the relevant IC constraint is the one that guarantees that it is not optimal to always announce L, we may show that it is optimal to tax savings.

4.5 Discussion

If, contrary to what was assumed in section 3, the government cannot directly control first period choices, it is still the case that it should intervene in the market by discouraging savings and encouraging education.

When the two instruments are used separately, this is exactly equivalent to taxation of savings and subsidization of education. However, when the two are used at the same time, the cross effects do not allow one to sign optimal tax rates.

Yet, it is possible to show that under the same conditions of the last two sections we have, $\tau\hat{h}_\tau + t\hat{s}_\tau > 0$ and $\tau\hat{h}_t + t\hat{s}_t < 0$, where variables with hats are 'compensated' choices.[11]

The caveat here is that we would want to interpret the two expressions as a statement about the encouragement of human capital investment and a discouragement of savings along the lines of Mirrlees (1976), however the association is not precise. The problem here is that, though $\hat{s}$ and $\hat{h}$ are 'compensated' savings and human capital investment, they need not be symmetric, i.e., $\hat{s}_\tau \neq \hat{h}_t$, in general.

Finally our prescription regarding the taxation of capital income is in contrast with the zero tax prescription found in most of the literature—see for example, Feldstein (1978) in a two-periods model and Chamley (1987) and Lucas (1990) in a Ramsey framework. It is, however, in line with the recent findings of Kocherlakota et. al. (2003).

Next we discuss a case where government is able to control investment in human capital but we maintain the assumption that savings are not observed.

The purpose of such an extension is to compare our results with one of the cases examined by Hamilton (1987). He considers an economy where the government may directly control investment in human capital but must rely on anonymous taxes on savings. He shows that with a multiplicative specification of human capital investment—which implies $w_{h\theta}(h, \theta) \geq$

[11] E.g., $\hat{s}_\tau \equiv s_\tau + \left(s_{y^H} + s_{y^L}\right) h$ measures the effect on savings of increasing taxes on human capital investment when this change is compensated with an increase of h units of income in each state of the world.

0—and under some other assumptions on labor supply, that, slightly reducing h is welfare increasing at the private optimal h and s. Moreover he points out the dependence of his result on the sign of the cross derivative of $w(\cdot)$. We strengthen his results by showing that, with fewer restrictions on labor supply functions and independently of the cross derivative of $w(\cdot)$, it is optimal to create a wedge between private marginal costs and benefits *at the optimum*.

4.6 Observable Education and Unobservable Savings

In this section we consider an economy where the government may directly control investment in human capital but must rely on anonymous taxes on savings.

Observable savings and unobservable education is not compatible with the framework we use, given that observing one implies observing the other. This is, of course, due to our simplifying assumptions regarding first period choices, more specifically the fact that, at $t = 0$, leisure is assumed to have no value. We set this aside while noting that the inclusion of a value for leisure would remove the inconsistency and solve the government's program

$$\max_{h,Y^H,Y^L,y^H,y^L} u(1-h-s^*) + \frac{1}{2}\sum_{i=H,L}\left[u(s^*+y^i) - \zeta\left(\frac{Y^i}{w(h,\theta^i)}\right)\right]$$

subject to

$$u(1-h-s^*) + \frac{1}{2}\sum_{i=H,L}\left[u(s^*+y^i) - \zeta\left(\frac{Y^i}{w(h,\theta^i)}\right)\right] \geq$$
$$u(1-h-s^o) + u(s^o+y^L) - \frac{1}{2}\sum_{i=H,L}\zeta\left(\frac{Y^L}{w(h,\theta^i)}\right), \qquad [\,\mu\,]$$

and

$$Y^H + Y^L \geq y^H + y^L, \qquad [\,\lambda\,]$$

where Lagrange multipliers are shown at the right of each constraint equation.

First order conditions for y^H and y^L are, respectively

$$u'\left(c^H\right) = \frac{\lambda}{1+\mu}, \quad \text{and} \quad u'(c^L) = \frac{\lambda}{1+\mu} - 2\frac{\mu}{1+\mu}u'(s^o+y^L),$$

while for Y^H and Y^L, respectively,

$$\frac{\zeta'\left(l^H\right)}{w(h,\theta^H)} = \frac{\lambda}{1+\mu}, \quad \text{and} \quad \frac{\zeta'\left(l^L\right)}{w(h,\theta^L)} = \frac{\lambda}{1+\mu} + \frac{\mu}{1+\mu}\left[\frac{\zeta'\left(l^L\right)}{w(h,\theta^L)} + \frac{\zeta'(l^{L|H})}{w(h,\theta^H)}\right].$$

Finally regarding h, we have,

$$2u'\left(1-h-s^{*}\right)=\sum_{i=H,L}\zeta'\left(l^{i}\right)l^{i}\frac{w_{h}(h,\theta^{i})}{w(h,\theta^{i})}+\mu\left[\zeta'\left(l^{H}\right)l^{H}-\zeta'\left(l^{L|H}\right)l^{L|H}\right]\frac{w_{h}(h,\theta^{H})}{w(h,\theta^{H})}.$$

Notice that the first order condition with respect to Y^{L} can be rewritten as

$$\frac{\zeta'\left(l^{L}\right)}{w(h,\theta^{L})}-u'(c^{L})=\frac{\mu}{1+\mu}\left\{\frac{1}{2}\left[\frac{\zeta'\left(l^{L}\right)}{w(h,\theta^{L})}+\frac{\zeta'(l^{L|H})}{w(h,\theta^{H})}\right]-u'(s^{o}+y^{L})\right\} \tag{18}$$

which parallels the one found in section 4. The difference is that human capital investments are controlled in (18), hence equal on and off the equilibrium path, whilst in section 4 they differ, in general.

High productivity agents will necessarily consume more—$(Y^{H},y^{H})>>(Y^{L},y^{L})$. The differences in marginal utilities will be proportional to $u'(s^{o}+y^{L})$, the marginal utility of consumption for those who opted for deviating. Notice also that:

$$\frac{\zeta'\left(l^{L}\right)}{w(h,\theta^{L})}>\frac{\zeta'\left(l^{H}\right)}{w(h,\theta^{H})}.$$

The marginal cost will be higher for the least productive.

So, the interpretation of the equation that determines the optimal quantity of education parallels that of section 3. Once again, a wedge between marginal cost and marginal benefit of education is induced by the government.

Equation (18) resembles (9). When choosing the optimal amount of education, the government should equalize its private cost, the foregone earnings, to its private benefit, the reduction in marginal cost of effort, plus one term associated to the incentive compatibility constraint. Noticing that $Y^{H}>Y^{L}$ one can easily see that this expression should be positive. Hence, we have the same situation as before: the government should distort the level of education to punish the deviators.

4.6.1 Taxing Savings

Assume that the government can tax savings as in section 4. In this case,

$$\frac{\partial\mathcal{L}}{\partial t}=(1+\mu)\left[u'(c^{H*})+u'(c^{L*})\right]s^{*}+\mu u'(c^{Lo})s^{o}+\lambda\left(s^{*}+t\frac{\partial s^{*}}{\partial t}\right)=0$$

The same procedure used in section 4 applies here. First we notice that

$$\frac{\partial\mathcal{L}}{\partial t}=\mu u'(c^{Lo})\left(s^{o}-s^{*}\right)+\lambda\frac{\partial\hat{s}^{*}}{\partial t}=0,$$

with

$$\frac{\partial \hat{s}^*}{\partial t} = \frac{\partial s^*}{\partial t} - \frac{1}{2}\left(\frac{\partial s^*}{\partial y^H} + \frac{\partial s^*}{\partial y^L}\right) s^* < 0$$

Hence, we only need to know the sign of $(s^o - s^*)$ to determine the sign of t.

But optimal savings on and off-equilibrium, s^* and s^o, can be determined by the following private first order conditions:

$$\begin{aligned} u'(1-h-s^*) &= \frac{1}{2}\left[u'(s^*+y^H) + u'(s^*+y^L)\right] \\ u'(1-h-s^o) &= u'(s^o+y^L) \end{aligned}$$

Hence, $s^o > s^*$. This shows that the optimal linear tax rates on savings are positive.

Finally, notice that an increase in h reduces savings (on and off the equilibrium path) and hence reducing tax revenues. This indirect effect must be accounted for by the government when deciding the optimal level of h. It does not change, however, the qualitative results.

5 Conclusion

We have analyzed optimal income taxation and educational policy in a model of human capital formation with wage uncertainty due to idiosyncratic shocks to productivity along the lines of Eaton and Rosen (1980) and Hamilton (1987).

Our model differs from these early contributions to the literature in two dimensions. First, we generalize the human capital investment technology by letting human capital either increase or decrease wage variance. Second, we allow the government to use a fully non-linear income tax schedule.

By generalizing investment technology we show that the prescriptions we derive are robust to the risk characteristics of human capital: government should encourage education and discourage savings. This is novel in the literature where it is explicitly recognized that the policy prescription usually found depend on the risk characteristics that are *imposed* on the nature of human capital investments.

Robustness is important for no conclusive evidence is yet available concerning risk characteristics of human capital. In our model, robustness is due to our prescriptions being driven by reasons that are different from the usual ones. The government should finance education and tax savings to punish those who intend to deviate and free-ride on the government's redistributive programs, claiming to be less productive than they really are. Because the pattern of choices driven by deviant behavior is independent of the risk characteristics of human capital investment, so are policy prescriptions.

We have also built a bridge between different ways of analyzing the problem according to specific assumptions about the informational structure and have shown that our results are also robust to different possibilities regarding the observability of investment choices. Specifically, wedges between marginal private costs and benefits are created with instruments that are feasible according to the informational structure of the problem. Signs of wedges, however, are invariant to these informational changes.

A Appendix

This appendix contains proofs of results stated in the main text.

Proof of Proposition 1. Because IC constraints (16) are satisfied we have

$$\begin{aligned}
&2u\left(1-s^{*}-h^{*}\right)+\sum_{i}\left\{u(s^{*}+y^{i})-\zeta\left(\frac{Y^{i}}{w(h^{*},\theta^{i})}\right)\right\}\geq\\
&2u\left(1-s^{o}-h^{o}\right)+\sum_{i}\left\{u(s^{o}+y^{\sigma^{o}(i)})-\zeta\left(\frac{Y^{\sigma^{o}(i)}}{w(h^{o},\theta^{i})}\right)\right\}\geq\\
&2u\left(1-s^{*}-h^{*}\right)+\sum_{i}\left\{u(s^{*}+y^{\sigma^{o}(i)})-\zeta\left(\frac{Y^{\sigma^{o}(i)}}{w(h^{o},\theta^{i})}\right)\right\}
\end{aligned}$$

where the last inequality is due to s^o and h^o being the optimal choices for alternative strategy o.

Take first strategy $\sigma^{o}\left(j\right)=L$, $j=H,L$. Then,

$$\begin{aligned}
\sum_{i}\left\{u(s^{*}+y^{i})-\zeta\left(\frac{Y^{i}}{w(h^{*},\theta^{i})}\right)\right\} &\geq u(s^{*}+y^{L})-\sum_{i}\zeta\left(\frac{Y^{L}}{w(h^{*},\theta^{i})}\right)\Rightarrow\\
u(s^{*}+y^{H})-\zeta\left(\frac{Y^{H}}{w(h^{*},\theta^{H})}\right) &\geq u(s^{*}+y^{L})-\zeta\left(\frac{Y^{L}}{w(h^{*},\theta^{H})}\right).
\end{aligned}$$

Therefore, one of the second period IC constraints is satisfied. To prove that the other is also satisfied we only need to consider the other alternative strategy $\sigma^{o}\left(j\right)=H$, $j=H,L$. An analogous procedure delivers the result. ■

Proof of Lemma 1. For notational simplicity, in this proof, we normalize $s^*=0$ and simplify notation by assuming $w\left(h^{*},\theta^{L}\right)=1$, $w\left(h^{*},\theta^{H}\right)=w>1$. Assume $\left(Y^{H},y^{H}\right)<<\left(Y^{L},y^{L}\right)$. Take any monotonic path $(y,Y):[0,1]\longrightarrow\left[y^{L},Y^{L}\right]\times\left[y^{H},Y^{H}\right]$ from $\left(y^{L},Y^{L}\right)$

to $\left(y^H, Y^H\right)$ and write

$$\begin{aligned} u(y^H) - \zeta\left(Y^H\right) - \left[u(y^L) - \zeta\left(Y^L\right)\right] &= \int_0^1 \left\{u'(y(t))y'(t) - \zeta'(Y(t))Y'(t)\right\} dt \\ &> \int_0^1 \left\{u'(y)y'(t) - \zeta'\left(\frac{Y(t)}{w}\right)\frac{Y'(t)}{w}\right\} dt \\ &= u(y^H) - \zeta\left(\frac{Y^H}{w}\right) - \left[u(y^L) - \zeta\left(\frac{Y^L}{w}\right)\right] \end{aligned}$$

If an allocation is implementable the left hand side is non-positive, which means that

$$u(y^L) - \zeta\left(\frac{Y^L}{w}\right) > u(y^H) - \zeta\left(\frac{Y^H}{w}\right).$$

Hence, the allocation is not implementable. ■

Proof of Proposition 2. Let s^o and h^o be the optimum level of savings and human capital investments for the strategy σ^o such that $\sigma^o(H) = L$ and $\sigma^o(L) = H$. Then, lemma 1, monotonicity and convexity of $\zeta(\cdot)$ guarantee that

$$\zeta\left(\frac{Y^H}{w(h^o,\theta^L)}\right) + \zeta\left(\frac{Y^L}{w(h^o,\theta^H)}\right) > \zeta\left(\frac{Y^H}{w(h^o,\theta^H)}\right) + \zeta\left(\frac{Y^L}{w(h^o,\theta^L)}\right)$$

which then implies,

$$\begin{aligned} &2u\left(1 - s^o - h^o\right) + 2u\left(y^H + s^o\right) - \left[\zeta\left(\frac{Y^H}{w(h^o,\theta^L)}\right) + \zeta\left(\frac{Y^L}{w(h^o,\theta^H)}\right)\right] < \\ &2u\left(1 - s^o - h^o\right) + 2u\left(y^H + s^o\right) - \left[\zeta\left(\frac{Y^H}{w(h^o,\theta^H)}\right) + \zeta\left(\frac{Y^L}{w(h^o,\theta^L)}\right)\right] \leq \\ &2u\left(1 - s^* - h^*\right) + 2u\left(y^H + s^*\right) - \left[\zeta\left(\frac{Y^H}{w(h^*,\theta^H)}\right) + \zeta\left(\frac{Y^L}{w(h^*,\theta^L)}\right)\right], \end{aligned}$$

where the last inequality is a consequence of (s^*, h^*) being optimal for the true-telling strategy. ■

Proof of Proposition 3. First, assume that the optimal allocation is such that $Y^H < y^H$ and $Y^L > y^L$. Then for any s and h the allocation is a mean preserving spread over the allocation that the agent would obtain in autarchy by choosing s and h and producing Y^H and Y^L conditional on having innate ability θ^H and θ^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

$$2u\left(1-s^{*}-h^{*}\right)+\sum_{i=H,L}\left[u\left(y^{i}+s^{*}\right)-\zeta\left(\frac{Y^{i}}{w(h^{*},\theta^{i})}\right)\right]=$$
$$2u\left(1-s^{o}-h^{o}\right)+2u\left(y^{H}+s^{o}\right)-\sum_{i=H,L}\zeta\left(\frac{Y^{H}}{w(h^{o},\theta^{i})}\right)\leq$$
$$2u\left(1-s^{o}-h^{o}\right)+2u\left(Y^{H}+s^{o}\right)-\sum_{i=H,L}\zeta\left(\frac{Y^{H}}{w(h^{o},\theta^{i})}\right),$$

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. ■

Proof of Proposition 4. Fix s^* and s^o and assume that $h^* \leq h^o$. The first order conditions for savings are

$$2u'\left(1-s^{*}-h^{*}\right)=u'\left(y^{H}+s^{*}\right)+u'\left(y^{L}+s^{*}\right) \tag{19}$$

and

$$u'\left(1-s^{o}-h^{o}\right)=u'\left(y^{L}+s^{o}\right) \tag{20}$$

with $y^L < y^H$. While the first order conditions for human capital investments imply, $2u'\left(1-s^{*}-h^{*}\right)=$

$$\left[\zeta'\left(\frac{Y^{H}}{w(h^{*},\theta^{H})}\right)\frac{Y^{H}\left(\partial w^{H}/\partial h\right)^{*}}{w(h^{*},\theta^{H})^{2}}+\zeta'\left(\frac{Y^{L}}{w(h^{*},\theta^{L})}\right)\frac{Y^{L}\left(\partial w^{L}/\partial h\right)^{*}}{w(h^{*},\theta^{L})^{2}}\right]\geq$$
$$\left[\zeta'\left(\frac{Y^{H}}{w(h^{o},\theta^{H})}\right)\frac{Y^{H}\left(\partial w^{H}/\partial h\right)^{o}}{w(h^{o},\theta^{H})^{2}}+\zeta'\left(\frac{Y^{L}}{w(h^{o},\theta^{L})}\right)\frac{Y^{L}\left(\partial w^{L}/\partial h\right)^{o}}{w(h^{o},\theta^{L})^{2}}\right]\geq$$
$$\left[\zeta'\left(\frac{Y^{L}}{w(h^{o},\theta^{H})}\right)\frac{Y^{L}\left(\partial w^{H}/\partial h\right)^{o}}{w(h^{o},\theta^{H})^{2}}+\zeta'\left(\frac{Y^{L}}{w(h^{o},\theta^{L})}\right)\frac{Y^{L}\left(\partial w^{L}/\partial h\right)^{o}}{w(h^{o},\theta^{L})^{2}}\right]=$$
$$2u'\left(1-s^{o}-h^{o}\right),$$

where $\left(\partial w^H/\partial h\right)^* \equiv \partial w(h^*,\theta^H)/\partial h$ and $\left(\partial w^H/\partial h\right)^o \equiv \partial w(h^o,\theta^H)/\partial h$.
The two equalities—which begin and end the expression—are simply the first order conditions with respect to h along and off the equilibrium path. The second inequality is a simple consequence of $Y^L < Y^H$. As for the first, it is due to the facts that: i) $h^* \leq h^o \Rightarrow w(h^*,\theta) < w(h^o,\theta)$ (monotonicity) and $(\partial w/\partial h)^o < (\partial w/\partial h)^*$ (concavity), and ii) convexity and monotonicity of $\zeta(\cdot)$.

Hence, $u'(1 - s^* - h^*) \geq u'(1 - s^o - h^o)$ for any s, $h^* \leq h^o$ which is not compatible with (19) and (20).

As for savings, assume $s^* > s^o$. Then,

$$u'\left(y^L + s^o\right) > \frac{1}{2}\left[u'\left(y^H + s^*\right) + u'\left(y^L + s^*\right)\right] = u'(1 - s^* - h^*) > u'(1 - s^o - h^o).$$

where the last inequality is due to $h^* > h^o$. This contradicts s^o and h^o satisfying the first order conditions. ∎

References

[1] **Albanesi, S. and C. Sleet** (2003) "Dynamic Optimal Taxation with Private Information," Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis Discussion Paper 140.

[2] **Anderson, D. and F. Anderberg** (2003) "Investments in human capital, wage uncertainty, and public policy" *Journal of Public Economics,* 87: 1521-1537.

[3] **Brett, C. and J. Weymark** (2000) "Financing education using optimal redistributive taxation." Paper presented at ISPE conference at CORE, Louvain.

[4] **Chamley, C.** (1987) "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives" *Econometrica*, 54, 607-622.

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

[6] **da Costa, C.** (2004). "Yet Another reason to tax Goods." MIMEO, EPGE-FGV.

[7] **da Costa, C. and I. Werning** (2001) "A Note on Pre-Committed Goods." MIMEO, University of Chicago.

[8] **Eaton, J. and H. Rosen** (1980) "Taxation, Human Capital, and Uncertainty." *American Economic Review*, 70, 705-715.

[9] **Feldstein, M.** (1978) "The Welfare Cost of Capital Income Taxation" *Journal of Political Economy*, 86, S29-S51.

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

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

[12] **Golosov, M. and A. Tsyvinski** (2003b) "Optimal Taxation with Endogenous Insurance". Mimeo, UCLA.

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

[14] **Hamilton, J.**(1987). "Optimal Wage and Income Taxation with Uncertainty." *International Economic Review*, 28, 373-388.

[15] **Hare, P. and D. Ulph** (1979) "On Education and Distribution" *Journal of Political Economy*, 87, S193-S212

[16] **Heckman, J.** (1976) "A Life-Cycle Model of Earnings, Learning, and Consumption" *Journal of Political Economy*, 84, S11-S44.

[17] **Hoff, K. and A. Lyon** (1995) "Non-leaky buckets: Optimal redistributive taxation and agency costs" *Journal of Public Economics,* 58: 365-390

[18] **Judd,** (1985), "Redistributive Taxation in a Simple Perfect Foresight Model", *Journal of Public Economics*, 28, 59-83.

[19] **Kocherlakota, N.** (2004) "Wedges and Taxes". *American Economic Review*, *Papers and Proceedings,* forthcoming.

[20] **Kocherlakota, N.** (2003) "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation" MIMEO. Stanford University.

[21] **Lucas, R.** (1990) "Supply-Side Economics" *Oxford Economic Papers*, 42, 293-316.

[22] **Mirrlees, J.** "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, 38, 175–208.

[23] **Mirrlees, J.** (1976), "Optimal Tax Theory: A Synthesis", *Journal of Public Economics*, 6, 327–358.

[24] **Naito, H.** (2004) "Endogenous human capital accumulation, comparative advantage and direct vs. indirect redistribution"*Journal of Public Economics*, forthcoming.