Optimal Income Taxation with Asset Accumulation

Árpád Ábrahám,[*] *Sebastian Koehne,*[†] and *Nicola Pavoni*[‡]

WORK IN PROGRESS

Please do not circulate without permission

November 2010

Abstract

Several frictions might prevent (or make undesirable) the full taxation of savings. Due to international capital mobility, for instance, the government may not have perfect control over agents' saving and consumption decisions. We show in this paper that a restricted ability to tax savings has important implications for the taxation of labor income. Specifically, when agents have preferences with convex absolute risk-aversion, we find that optimal marginal tax rates on labor income become more regressive when savings cannot be fully taxed.

We derive this result in a two-period model of social insurance. By exerting effort, agents can influence their labor income realizations. Moreover, agents can accumulate a risk-free bond. We show that the trade-offs in designing optimal effort incentives depend crucially on the tax rate on bond returns. Through this channel, the ability of taxing savings interacts with the structure of optimal labor income taxes.

Keywords: Optimal Income Taxation, Asset Accumulation, Progressivity.

JEL: D82, D86, E21, H21.

1 Introduction

The progressivity of the income tax code is a central question in the economic literature and the public debate. While we observe progressive tax systems in most developed countries nowadays, theoretical insights on progressivity are rather limited. Previous research has examined how the optimal degree of progressivity depends on the skill distribution (Mirrlees 1971), the welfare criterion (Sadka 1976), and earnings elasticities (Saez 2001).

[*]European University Institute, Florence. E-mail address: `Arpad.Abraham@EUI.eu`

[†]University of Mannheim. E-mail address: `skoehne@mail.uni-mannheim.de`

[‡]EUI, IFS, IGIER, and CEPR. E-mail address: `nicola.pavoni@unibocconi.it`

The existing approaches to income tax progressivity have largely focused on models where labor is the only source of income. In the present paper, we argue that the optimal shape of labor income taxes cannot be determined in isolation from the tax code on capital income. Specifically, we show that taxing capital has an important effect on the optimal progressivity of labor income taxes: When capital is taxed, the optimal tax on labor income becomes more progressive.

We derive this result in a two-period model of social insurance. A continuum of ex-ante identical agents influence their labor incomes by exerting effort. Labor income realizations are not perfectly controllable, which creates a moral hazard problem. The social planner faces a trade-off between insuring the agents against idiosyncratic income uncertainty on the one hand and the associated disincentive effects on the other hand. In addition, agents have access to a risk-free bond, which gives them a limited means for self-insurance.

There is an efficiency reason for taxing bond returns in our model: The bond provides insurance against the realization of labor income and thereby reduces the incentives to exert effort. Consequently, the planner faces an additional constraint when determining the optimal social insurance scheme, since she has to satisfy the agent's Euler equation. However, if the bond is appropriately taxed, the constraint becomes non-binding, which crucially changes the cross-sectional structure of the planner's costs of allocating utility. We examine this effect in detail and relate it to the shape of the agent's absolute risk aversion. If absolute risk-aversion is convex, we find that optimal consumption moves in a more concave way with labor income when bond returns are (optimally) taxed. Equivalently, the marginal tax on labor income becomes more progressive.

To the best of our knowledge, this is the first paper that examines how capital taxation affects the optimal progressivity of the labor income tax code. Recent work on dynamic Mirrleesian taxation models has highlighted a somewhat complementary question. Kocherlakota (2005), Albanesi and Sleet (2006), and Golosov and Tsyvinski (2006) focus on the optimal taxation of capital, given nonlinear taxation of income. In that literature, the reason for capital taxation is similar to our model and stems from disincentive effects associated with the accumulation of wealth.[1] While the Mirrlees (1971) framework focuses on redistribution in a population with heterogeneous skills, our approach highlights the social insurance (or ex-post redistribution) aspect of income taxation. In spirit, our model is therefore closer to the works by Varian (1980) and Eaton and Rosen (1980).

An entirely different link between labor taxation and capital taxation is explored by Conesa, Kitao and Krueger (2009). Using a life-cycle model with time-varying labor supply elasticities, they

[1]However, the tax on capital takes a much simpler form in our model: a linear tax on the aggregate trade of the bond is sufficient to implement the second best; compare Gottardi and Pavoni (2010). Notice, in particular, that this tax can be implemented in an anonymous asset market where the planner only observes the aggregate trading volume.

argue that capital taxes and progressive labor income taxes are two alternative ways of mimicking age-dependent taxation. They then use quantitative methods to determine the right relation between the two instruments. Interestingly, in the present environment we obtain a very different conclusion. While Conesa, Kitao and Krueger (2009) argue that capital taxes and progressive labor taxes are essentially substitutes, in our model they are complements. Laroque (2010) derives - analytically - a similar substitutability between income and capital taxes restricting labor taxation to be nonlinear but homogenous across age groups. In both these cases, the substitutability arises since exogenous taxes are in general imperfect instruments to perform redistribution, and wealth is typically positively correlated to skill, hence the optimality of compensating the imperfect distribution coming from income taxation with capital taxation.

The paper proceeds as follows: Section 2 describes the setup of the model. Section 3 presents the main result of the paper: Capital taxation increases the progressivity of the optimal income tax code. In Section 4, we explore alternative concepts of concavity/progressivity. Section 5 explores the quantitative importance of our results, while Section 6 concludes.

2 Model

Consider a benevolent social planner (the principal) whose objective is to maximize the welfare of the citizens. The (small open) economy consists of a continuum of ex-ante identical agents who live for two periods, $t = 0, 1$, and can influence their date-1 income realizations by working hard or shirking. The planner offers a tax/transfer system to insure them against idiosyncratic risk and provide them appropriate incentives for working hard. The planner's budget must be (intertemporally) balanced.

Preferences The agent derives utility from consumption $c_t \geq \underline{c} \geq -\infty$ and effort $\infty \geq \bar{e} \geq e_t \geq 0$ according to: $u(c_t) - v(e_t)$, where both u and v are strictly increasing and twice continuously differentiable functions, and u is strictly concave whereas v is convex. We normalize $v(0) = 0$. The agent's discount factor is denoted by $\beta > 0$.

Production and endowments At date $t = 0$, the agent has a *fixed* endowment y_0.[2] At date $t = 1$, the agent has a stochastic income $y \in Y := [\underline{y}, \overline{y}]$. The realization of y is publicly observable, while the probability distribution over Y is affected by the agent's unobservable effort level e_0 that is exerted at $t = 0$. The probability density of this distribution is given by the smooth

[2] We could allow for *observable* ex-ante heterogeneity on y_0 (and f depending on y_0) with no changes in our results.

function $f(y, e_0)$. As in most of the the optimal contracting literature, we assume *full support*, that is $f(y, e_0) > 0$ for all y, e_0. There is no production or any other action at $t \geq 2$.

Markets At each date, the agent can buy or (short)-sell a risk-free bond b_t which costs $q \geq 0$ consumption units today and pays one unit of consumption tomorrow. The agent has no access to any insurance market other than that delivered by the planner. The planner can impose a linear tax τ^k on the price of the bond.[3] Therefore, the net price of the bond is $\tilde{q} = (1 + \tau^k)q$.

There are two ways of motivating the linearity assumption of the tax τ^k. First, the planner only needs to observe the aggregate trade of the bond in order to implement such a tax. Therefore, the tax is feasible even in an anonymous market where individual asset decisions and consumption levels are private information. Second, linearity is in fact without loss of generality in the present model, since the planner is able to generate the second best with such a tax (see Proposition 1).

Given the structure of the problem, the agent will never be able to borrow at $t = 1$, hence we have $b_1 \geq 0$. Monotonicity of preferences guarantees that the agent will not want to leave any positive amount of assets at date 1 either. So, $b_1 = 0$ for all states y. Similarly, since v is strictly increasing, $e_1 = 0$ for all states y.

Transfer Schemes A *social contract* $\mathcal{W} := (\boldsymbol{T}, \tau^k, e_0, b_0)$ consists of a tax/transfer scheme $\boldsymbol{T}$, a tax rate on the bond τ^k, and recommendations (e_0, b_0). The tax/transfer scheme $\boldsymbol{T}:= (T_0, T(\cdot))$ has two components: T_0 denotes the transfer the individual receives in period $t = 0$, and $T(y)$, $y \in Y$, denotes the transfer the individual receives in period $t = 1$ conditional on income realization y.

Given a contract $\mathcal{W}$, the agent's utility is

$$U(e_0, b_0; \boldsymbol{T}, \tau^k) := u(y_0 + T_0 - (1 + \tau^k)qb_0) - v(e_0) + \beta \int_{\underline{y}}^{\overline{y}} u(y + T(y) + b_0) f(y, e_0)\, dy.$$

To guarantee solvency of the agent for every contingency, we impose the 'natural' borrowing limit: $b_0 \geq \underline{c} - \inf_y \{y + T(y)\}$.

The social planner faces the same credit market as the agent, therefore her discount rate is q. The planner's expenditures are

$$T_0 + q \int_{\underline{y}}^{\overline{y}} T(y) f(y, e_0)\, dy - \tau^k q b_0 + G,$$

where G denotes government consumption.

[3] A tax on the bond price is equivalent to a tax on the return in our model.

Efficiency An *optimal social contract* is a contract that maximizes ex-ante welfare

$$\max_{\mathcal{W}} \quad U(e_0, b_0; \boldsymbol{T}, \tau^k) \tag{1}$$

subject to the planner's *budget constraint*

$$-T_0 - q \int_{\underline{y}}^{\overline{y}} T(y) f(y, e_0)\, dy + \tau^k q b_0 - G \geq 0 \tag{2}$$

and the *incentive compatibility constraint*

$$(e_0, b_0) \in \arg \left\{ \max_{e,b} \ U(e, b; \boldsymbol{T}, \tau^k) \text{ s.t. } e \geq 0,\ y_0 + T_0 - \underline{c} \geq \tilde{q} b \geq -\tilde{q} \inf_{y} \{ y + T(y) - \underline{c} \} \right\}. \tag{3}$$

Note that there is indeterminacy in the contract between T_0 and b_0. The planner can implement the same allocation with a contract

$$\Big(T_0, T(\cdot), \tau^k, e_0, b_0 \Big)$$

and with a contract

$$\Big(T_0 - \tilde{q}\varepsilon, T(\cdot) + \varepsilon, \tau^k, e_0, b_0 - \varepsilon \Big).$$

In other words, since the planner and the agent face the same credit market, there is a continuum of optimal contracts. Throughout this paper, without loss of generality, we will study the one specific optimal contract that implements $b_0 = 0$. Because of these observations, we will sometimes refer to the combination of e_0, τ^k, and $\boldsymbol{c} = (c_0, c(\cdot))$, with $c_0 := y_0 + T_0$, $c(y) := y + T(y)$, $y \in Y$, as a contract.

First-order approach Throughout this paper, we assume that the first-order approach (FOA) is justified. Hence, we can replace the incentive constraint (3) by the first-order conditions of the agent's maximization problem with respect to e_0 and b_0. Sufficient conditions for the validity of the FOA in this setup are given in Abraham, Koehne, and Pavoni (2010). Specifically, the FOA is valid if the agent has nonincreasing absolute risk aversion and the cumulative distribution function of income is log-convex in effort.[4]

Using the normalization $b_0 = 0$ and the notation $\tilde{q} = (1 + \tau^k)q$, $c_0 = y_0 + T_0$, $c(y) = y + T(y)$, $y \in Y$, we can thus rewrite the planner's problem as

$$\max_{\mathbf{c}, \tilde{q}, e_0} \quad u(c_0) - v(e_0) + \beta \int_{\underline{y}}^{\overline{y}} u(c(y)) f(y, e_0)\, dy \tag{4}$$

[4] As argued by Abraham, Koehne, and Pavoni (2010), both conditions have quite a broad empirical support. First, virtually all estimations for u reveal NIARA; see Guiso *et al.* (2001) for example. The condition on the distribution function essentially restricts the agent's Frisch elasticity of labor supply. This restriction is satisfied as long as the Frisch elasticity is smaller than unity. In fact, most empirical studies find values for this elasticity between 0 and 0.5; see Domeij and Floden (2006), for instance.

subject to $c_0 \geq \underline{c}$, $c(y) \geq \underline{c}$, $e_0 \geq 0$, the planner's budget constraint

$$y_0 - c_0 + q \int_{\underline{y}}^{\overline{y}} (y - c(y)) f(y, e_0)\, dy - G \geq 0 \tag{5}$$

and the first-order incentive conditions

$$-v'(e_0) + \beta \int_{\underline{y}}^{\overline{y}} u(c(y)) f_e(y, e_0)\, dy \;\; \geq \;\; 0 \tag{6}$$

$$\tilde{q} u'(c_0) - \beta \int_{\underline{y}}^{\overline{y}} u'(c(y)) f(y, e_0)\, dy \;\; \geq \;\; 0. \tag{7}$$

Under mild assumptions, the optimal contract is interior: $c_0 > \underline{c}$, $c(y) > \underline{c}$, $e_0 > 0$.[5] In this case, using λ, μ and ξ as the (nonnegative) Lagrange multipliers associated with the constraints (5), (6), (7), respectively, the first-order conditions of the Lagrangian with respect to consumption are

$$\frac{\lambda q}{\beta u'(c(y))} \;\; = \;\; 1 + \mu \frac{f_e(y, e_0)}{f(y, e_0)} + \xi a(c(y)), \quad y \in [\underline{y}, \overline{y}], \tag{8}$$

$$\frac{\lambda}{u'(c_0)} \;\; = \;\; 1 - \xi \tilde{q} a(c_0), \tag{9}$$

where $a(c) := -u''(c)/u'(c)$ denotes the agent's absolute risk aversion.

Preliminary characterization of optimal contracts Because of the incentive problem, it is efficient to impose a positive tax on savings in our model; compare Gottardi and Pavoni (2010).

Proposition 1 Assume that the FOA is justified and that the optimal contract is interior. Then the tax on savings is positive: $\tau^k > 0$. Moreover, equations (8) and (9) characterizing the consumption scheme are satisfied with $\xi = 0$.

Proof. See Gottardi and Pavoni (2010). **Q.E.D.**

The above result is intuitive. It is efficient to tax the bond, because saving provides insurance against the incentive scheme. By appropriately reducing the rate of return, however, the planner can control the agent's intertemporal decision and therefore circumvent the (first-order) incentive constraint for saving.[6] Consequently, when savings can be taxed, condition (8) takes the form that is familiar from dynamic moral hazard models without asset accumulation:

$$\frac{\lambda q}{\beta u'(c(y))} = 1 + \mu \frac{f_e(y, e_0)}{f(y, e_0)}, \quad y \in [\underline{y}, \overline{y}].$$

[5] A sufficient condition for interiority is $\lim_{c \to \underline{c}} u'(c) = \infty$, $v'(0) = 0$.

[6] The validity of the first-order approach is crucial here, since it allows to characterize the agent's saving decision exclusively based on the rate of return.

Since we are interested in understanding how the shape of optimal consumption depends on the possibility of taxing savings, we contrast the optimal contract when τ^k is a choice variable for the planner with the optimal contract when τ^k is restricted to zero (i.e., $\tilde{q} = q$).

Proposition 2 Consider the above problem with τ^k restricted to zero. Assume that the FOA is justified and that the optimal contract is interior. Then equations (8) and (9) characterizing the consumption scheme are satisfied with $\xi > 0$.

Proof. Recall that, from the Kuhn-Tucker theorem, $\xi \geq 0$. If $\xi > 0$, we are done. If $\xi = 0$, then the first-order conditions of the Lagrangian read

$$\begin{aligned} \frac{\lambda}{u'(c_0)} &= 1, \\ \frac{\lambda q}{\beta u'(c(y))} &= 1 + \mu \frac{f_e(y, e_0)}{f(y, e_0)}, \quad y \in [\underline{y}, \overline{y}]. \end{aligned}$$

Since $f(y, e)$ is a density, integration of the last line yields

$$\int_{\underline{y}}^{\overline{y}} \frac{\lambda q}{\beta u'(c(y))} f(y, e_0)\, dy = 1.$$

As a consequence, we obtain

$$\frac{\lambda}{u'(c_0)} = \int_{\underline{y}}^{\overline{y}} \frac{\lambda q}{\beta u'(c(y))} f(y, e_0)\, dy \geq \frac{\lambda q}{\beta \int_{\underline{y}}^{\overline{y}} u'(c(y)) f(y, e_0)\, dy},$$

where the inequality follows from Jensen's inequality. The inequality is in fact strict, since the agent cannot be fully insured when effort is interior. Hence, we conclude

$$\lambda \beta \int_{\underline{y}}^{\overline{y}} u'(c(y)) f(y, e_0)\, dy > \lambda q u'(c_0).$$

For $\tau^k = 0$, we have $q = \tilde{q}$, however. Therefore, the above inequality is incompatible with the agent's Euler equation (7). This shows that ξ cannot be zero. **Q.E.D.**

3 Absolute Progressivity and Linear Likelihoods

We are interested in the shape of the optimal tax/transfer scheme $\boldsymbol{T}$. Clearly, this shape is closely related to the curvature of consumption $c(y) = y + T(y)$. Recall that we can fix $b_0 = 0$ without loss of generality.

Definition 1 We say that the transfer scheme $\boldsymbol{T}$ is *progressive (regressive)* if $c'(y)$ is decreasing (increasing) in y. We call $\boldsymbol{T}$ *proportional* if $c'(y)$ is constant in y.

This definition implies that whenever consumption is a concave (convex) function of income we have a progressive (regressive) tax system supporting it. In terms of the taxes and transfers $T(y)$, in a progressive system taxes ($T(y) < 0$) are increasing faster than income does. At the same time, for the states when the agent is receiving a transfer ($T(y) > 0$), transfers are increasing slower than income is decreasing. The opposite happens when we have a regressive scheme. Intuitively, if the scheme is progressive, incentives are provided more by imposing 'large penalties' for low income realizations, since consumption decreases relatively quickly when income decreases. Regressive schemes, by contrast, put more emphasis on rewards for high income levels than punishments for low income levels.

The next proposition provides conditions for progressivity and regressivity of the optimal scheme.

Proposition 3 (Sufficient conditions for progressivity/regressivity) Assume that the FOA is justified and that the optimal contract is interior.

(i) If the likelihood ratio function $l(y,e) := \frac{f_e(y,e)}{f(y,e)}$ is concave in y and $\frac{1}{u'(c)}$ is convex in c, then $\boldsymbol{T}$ is progressive. If, in addition, absolute risk aversion $a(c)$ is decreasing and concave,[7] then this result continues to hold when τ^k is restricted to zero.

(ii) On the other hand, if $l(y,e)$ is convex in y and $\frac{1}{u'(c)}$ is concave in c, then $\boldsymbol{T}$ is regressive. If, in addition, absolute risk aversion $a(c)$ is decreasing and convex, then this result continues to hold when τ^k is restricted to zero.

Proof. We only show (i), since statement (ii) can be seen analogously. Define

$$g(c) := \frac{\lambda q}{\beta u'(c)} - \xi a(c).$$

Notice that $\frac{1}{u'(\cdot)}$ is always increasing. Therefore, if $\frac{1}{u'(\cdot)}$ is convex and $\xi = 0$ (or $\xi > 0$ and $a(\cdot)$ decreasing and concave), then $g(\cdot)$ is increasing and convex. Given the validity of the FOA, Proposition 1 (Proposition 2) shows that optimal consumption is defined as follows:

$$g(c(y)) = 1 + \mu\, l(y, e_0),$$

where, by assumption, the right-hand side is a positive affine transformation of a concave function. By applying the inverse function of $g(\cdot)$ to both sides, we see that $c(\cdot)$ is concave since it is an increasing and concave transformation of a concave function. **Q.E.D.**

[7] Notice that absolute risk aversion is bounded below by zero. Therefore, the function $a(\cdot)$ can only be decreasing and concave over $[0,\infty)$ if it is constant.

Note that in the previous proposition, since the function g is increasing, consumption is increasing as long as the likelihood ratio function $l(y,e)$ is increasing in y.

Proposition 3 implies that CARA utilities with concave likelihood ratios lead to progressive schemes, no matter whether savings are taxed or not.[8] When savings are taxed, progressive schemes are also induced by concave likelihood ratios and CRRA utilities with $\sigma \geq 1$, since $\frac{1}{u'(c)} = c^{\sigma}$ is convex in this case. For logarithmic utility with linear likelihood ratios we obtain a scheme that is *proportional*, since $\frac{1}{u'(c)} = c$ is both concave and convex. Interestingly, when savings are not taxed, the scheme becomes *regressive* in this case (since absolute risk aversion $a(c) = \frac{1}{c}$ is convex).[9] This particular finding sheds light on a more general pattern under convex absolute risk aversion: when savings are taxed, the allocation has a 'more concave' relationship between income and consumption. In other words, taxing savings calls for more progressivity in the income tax/transfer system. The next result formalizes this insight.

Proposition 4 (Concavity) Assume that the FOA is justified. Let $\boldsymbol{c}$ be an interior, monotonic optimal consumption scheme for the general model and let $\hat{\boldsymbol{c}}$ be an interior, monotonic optimal consumption scheme for the model when τ^k is restricted to zero, both implementing effort level e_0. Moreover, assume that u has convex absolute risk aversion and that the likelihood ratio $l(y,e_0)$ is linear in y. Under these conditions, if $\hat{c}$ changes with y in a concave way, then c does as well.

Proof. Given validity of the FOA, $c(y)$ and $\hat{c}(y)$ are defined as follows (see Propositions 1 and 2):

$$g_{\lambda}(c(y)) = 1 + \mu\, l(y,e_0), \text{ where } g_{\lambda}(c) := \frac{\lambda q}{\beta u'(c)}, \tag{10}$$

$$\hat{g}_{\hat{\lambda},\hat{\xi}}(\hat{c}(y)) = 1 + \hat{\mu}\, l(y,e_0), \text{ where } \hat{g}_{\hat{\lambda},\hat{\xi}}(c) := \frac{\hat{\lambda} q}{\beta u'(c)} - \hat{\xi} a(c), \text{ with } \hat{\xi} > 0. \tag{11}$$

Since $l(y,e)$ is linear in y by assumption, concavity of $\hat{c}$ is equivalent to convexity of $\hat{g}_{\hat{\lambda},\hat{\xi}}$. Moreover, since $a(c)$ is convex in c by assumption, convexity of $\hat{g}_{\hat{\lambda},\hat{\xi}}$ implies convexity of $g_{\lambda} = \frac{\lambda}{\hat{\lambda}}\left(\hat{g}_{\hat{\lambda},\hat{\xi}} + \hat{\xi} a\right)$. Finally, notice that convexity of g_{λ} is equivalent to concavity of c, since $l(y,e)$ is linear in y. **Q.E.D.**

In order to obtain a clearer intuition of this result, we further examine condition (8), namely

$$\frac{\lambda q}{\beta u'(c(y))} = 1 + \mu \frac{f_e(y,e_0)}{f(y,e_0)} + \xi a(c(y)).$$

[8] Other cases where the tax on savings does not affect regressivity/progressivity are when a has the same shape as $\frac{1}{u'}$ (quadratic utility) and when a is linear (and hence increasing).

[9] More precisely, consumption is characterized by $\frac{\lambda q}{\beta} c(y) - \xi \frac{1}{c(y)} = 1 + \mu\, l(y,e)$ in this case. Since the left-hand side is concave in c and the right-hand side is linear in y, the consumption scheme $c(y)$ must be convex in y.

This expression equates the discounted present value (normalized by $f(y, e_0)$) of the costs and benefits of increasing the agent's utility by one unit in state y. The increase in utility costs $\frac{q}{\beta u'(c(y))}$ units in consumption terms. Multiplied by the shadow price of resources λ, we obtain the left-hand side of the above expression. In terms of benefits, first of all, since the agent's utility is increased by one unit, there is a return of 1. Furthermore, increasing the agent's utility also relaxes the incentive constraint for effort, generating a return of $\mu \frac{f_e(y,e_0)}{f(y,e_0)}$.[10] Finally, by increasing $u(c(y))$ the planner alleviates the saving motives of the agent. This gain, measured by $\xi a(c(y))$, depends crucially on whether savings are taxed or not. When savings are appropriately taxed, we have $\xi = 0$ and this gain vanishes. Intuitively, by controlling the net price of the bond, the planner is able to circumvent the incentive constraint for saving. However, when a tax on savings is ruled out, this constraint is binding and we have $\xi > 0$. Under convex absolute risk aversion, the gain $\xi a(c(y))$ is convex. This implies that, ceteris paribus, the benefits of increasing the agent's utility change in a more convex way with income. As a consequence, at the optimal contract the costs of increasing the agent's utility must also change in a more convex way with income, hence consumption becomes more convex in y in this case.

A closely related intuition for equation (8) can be obtained by rewriting it as follows:

$$\frac{\lambda q}{\beta u'(c(y))} - \xi a(c(y)) = 1 + \mu \frac{f_e(y, e_0)}{f(y, e_0)}. \tag{12}$$

On the right-hand side, we have the (rescaled) likelihood ratio. As in the static moral hazard problem, this function governs the allocation of utility across income states y. The only change compared to the static problem is the term $\xi a(c(y))$ on the left-hand side. This term stems from the agent's Euler equation and modifies the planner's costs of allocating utility over states. In the static model, allocating utility only generates a direct resource cost to the planner. This cost, captured by the discounted inverse marginal utility, is also present here. In addition, allocating utility to state y affects the intertemporal structure of the consumption scheme, which creates an additional cost due to the agent's Euler equation.

4 General Results on Progressivity

Since at least Holmstrom (1979), it is well understood that consumption patterns under moral hazard are crucially influenced by the shape of the likelihood ratio function $l(\cdot, e)$. Stated in more negative terms, one can always find functions $l(\cdot, e)$ so that the shape of consumption is almost arbitrary. To make the impact of the savings tax on the shape of optimal consumption easier to

[10] Of course, if the increase in consumption is done in a state with a negative likelihood ratio, this represents a cost since the incentive constraint is in fact tightened.

observe, we have therefore normalized the curvature of the likelihood ratio by assuming linearity in Proposition 4.

In this section, we study how the savings tax changes the curvature of the consumption scheme for arbitrary likelihood ratio functions. As usual, we assume that the FOA is justified and that $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ are interior, monotonic optimal contracts for the general model and the model without the savings tax, respectively, implementing the same effort level e_0.

Probably the most well known ranking in terms of concavity in economics is that dictated by concave transformations (e.g., Gollier 2001).

Definition 2 We say that f_1 is a concave (convex) transformation of f_2 if there is an increasing and concave (convex) function v such that $f_1 = v \circ f_2$.

Proposition 5 Let $\mathbf{c}$ and $\hat{\mathbf{c}}$ be as in Proposition 4. Assume that u has convex absolute risk aversion. Then, if $\hat{c}$ is a concave transformation of l, then c is a concave transformation of l. Conversely, if c is a convex transformation of l, then $\hat{c}$ has the same property.

Proof. Recall that we have

$$g_\lambda\left(c(y)\right) = 1 + \mu\, l\left(y, e_0\right), \tag{13}$$

$$\hat{g}_{\hat{\lambda},\hat{\xi}}\left(\hat{c}(y)\right) = 1 + \hat{\mu}\, l\left(y, e_0\right), \tag{14}$$

where the functions g_λ and $\hat{g}_{\hat{\lambda},\hat{\xi}}$ are defined as in (10) and (11), respectively. First, suppose that $\hat{c}$ is a concave transformation of l. Since the right-hand side of (14) is a positive affine transformation of l, this implies that $\hat{g}_{\hat{\lambda},\hat{\xi}}$ is convex. Now, notice that convexity of $\hat{g}_{\hat{\lambda},\hat{\xi}}$ implies that $g_\lambda\left(c\right) = \frac{\lambda}{\hat{\lambda}}\left(\hat{g}_{\hat{\lambda},\hat{\xi}}(c) + \hat{\xi}a(c)\right)$ is convex as well (since $a(c)$ is convex by assumption). Hence, using (13), we see that c is a concave transformation of l.

Conversely, suppose that c is a convex transformation of l. Using (13), we see that g_λ is then concave. Convexity of $a(c)$ implies that $\hat{g}_{\hat{\lambda},\hat{\xi}}$ is then also concave, which shows that $\hat{c}$ is a convex transformation of l. **Q.E.D.**

The previous finding induces an ordering that has the flavor of c being 'more concave' than $\hat{c}$. Note that this result generalizes Proposition 4 to arbitrary shapes of the likelihood ratio function l. As a drawback, we can rank the curvature of c and $\hat{c}$ only when, for example, c is more concave than l. We will now reduce the set of possible utility functions to facilitate such comparisons.

Let us consider the class of HARA (or linear risk tolerance) utility functions, namely

$$u\left(c\right) = \rho\left(\eta + \frac{c}{\gamma}\right)^{1-\gamma}$$

$$\text{with } \rho\frac{1-\gamma}{\gamma} > 0, \text{ and } \eta + \frac{c}{\gamma} > 0.$$

For this class, we have $a(c) = \left(\eta + \frac{c}{\gamma}\right)^{-1}$. Hence, absolute risk aversion is convex. Special cases of the HARA class are CRRA, CARA, and quadratic utility (e.g., see Gollier 2001).

Lemma Given a utility function $u : C \to \mathbb{R}$, consider the two functions defined as follows:

$$\begin{aligned} g_\lambda(c) &:= \frac{\lambda q}{\beta u'(c)}, \\ \hat{g}_{\hat{\lambda},\hat{\xi}}(c) &:= \frac{\hat{\lambda} q}{\beta u'(c)} - \hat{\xi} a(c). \end{aligned}$$

Then, if u belongs to the HARA class with $\gamma \geq -1$, then $\hat{g}_{\hat{\lambda},\hat{\xi}}$ is a concave transformation of g_λ for all $\hat{\lambda}, \hat{\xi} \geq 0$, $\lambda > 0$.

Proof. If u belongs to the HARA class, we obtain

$$\hat{g}_{\hat{\lambda},\hat{\xi}}(c) = \frac{\hat{\lambda}}{\lambda} g_\lambda(c) - \hat{\xi} a(c) = \frac{\hat{\lambda}}{\lambda} g_\lambda(c) - \hat{\xi} \lambda^{\frac{1}{\gamma}} \kappa \left(g_\lambda(c)\right)^{-\frac{1}{\gamma}}, \text{ with } \kappa = \left[\frac{\gamma q}{\beta\rho(1-\gamma)}\right]^{\frac{1}{\gamma}} > 0.$$

In other words, we have

$$\hat{g}_{\hat{\lambda},\hat{\xi}}(c) = h\left(g_\lambda(c)\right), \text{ where } h(g) = \frac{\hat{\lambda}}{\lambda} g - \hat{\xi} \lambda^{\frac{1}{\gamma}} \kappa\, g^{-\frac{1}{\gamma}}.$$

The second derivative of h with respect to g is $-\frac{\hat{\xi}\lambda^{\frac{1}{\gamma}}\kappa}{\gamma}\left(\frac{1}{\gamma}+1\right) g^{-\frac{1}{\gamma}-2}$, which is negative whenever $\gamma \geq -1$. **Q.E.D.**

The restriction $\gamma \geq -1$ in the above result is innocuous and allows for all HARA functions with nonincreasing absolute risk aversion as well as quadratic utility, for instance. To state the consequences of this Lemma, we introduce the concept of G-convexity (e.g., see Avriel et al., 1988), which is widely used in optimization. A function f is G-convex if once we transform f with G we get a convex function. More formally:

Definition 3 Let f be a function and G an increasing function mapping from the image of f to the real numbers. The function f is called *G-convex (G-concave)* if $G \circ f$ is a convex (concave) function.

This concept generalizes the standard notion of convexity. It is easy to see that a function f is convex if and only if it is G-convex for any increasing affine function G. Moreover, it can be shown that if G is concave and f is G-convex then f must be convex, but the converse is false.[11]

[11] For example, suppose $f(x) = x^2$ and $G(\cdot) = \log(\cdot)$, then $G(f(x)) = 2\log(x)$, which is obviously not convex.

Proposition 6 Assume u belongs to the HARA class with $\gamma \geq -1$. Then c is g_λ-convex (g_λ-concave) if and only if $\hat{c}$ is $\hat{g}_{\hat{\lambda},\hat{\xi}}$-convex ($\hat{g}_{\hat{\lambda},\hat{\xi}}$-concave).[12]

Proof. Recall that consumption is determined as follows:

$$\begin{aligned} g_\lambda\left(c(y)\right) &= 1 + \mu\, l\left(y, e_0\right), \\ \hat{g}_{\hat{\lambda},\hat{\xi}}\left(\hat{c}(y)\right) &= 1 + \hat{\mu}\, l\left(y, e_0\right). \end{aligned}$$

As a consequence, we can relate the two consumption functions as follows:

$$\frac{1}{\mu}\Big(g_\lambda\left(c(y)\right) - 1\Big) = \frac{1}{\hat{\mu}}\Big(\hat{g}_{\hat{\lambda},\hat{\xi}}\left(\hat{c}(y)\right) - 1\Big). \tag{15}$$

Now the result follows from the simple fact that convexity/concavity is preserved under positive affine transformations. **Q.E.D.**

Corollary If $\hat{c}$ is g_λ-concave then c is g_λ-concave. Conversely, if c is g_λ-convex then $\hat{c}$ is g_λ-convex.

Proof. Let $\hat{c}$ be g_λ-concave. By the Lemma, we have $\hat{g}_{\hat{\lambda},\hat{\xi}} = h \circ g_\lambda$ for some increasing and concave function h. Hence, when $\hat{c}$ is g_λ-concave, then $\hat{c}$ must also be $\hat{g}_{\hat{\lambda},\hat{\xi}}$-concave. Now Proposition 6 implies that c is g_λ-concave

To verify the second statement, let c be g_λ-convex. From Proposition 6, we see that $\hat{c}$ is $\hat{g}_{\hat{\lambda},\hat{\xi}}$-convex, i.e., $\hat{g}_{\hat{\lambda},\hat{\xi}} \circ \hat{c}$ is convex. By the Lemma, we have $\hat{g}_{\hat{\lambda},\hat{\xi}} = h \circ g_\lambda$ for some increasing and concave function h. Since the inverse of h must be convex, we conclude that $g_\lambda \circ \hat{c} = h^{-1} \circ \hat{g}_{\hat{\lambda},\hat{\xi}} \circ \hat{c}$ is convex. **Q.E.D.**

The corollary shows that whenever $\hat{c}$ satisfies the g_λ-concavity property, then c satisfies this property. In this sense, we note again that c is 'more concave' than $\hat{c}$.

Finally, it appears natural to ask whether the concavity of c and $\hat{c}$ can also be ranked according to the concavity notion of Definition 2. In other words, can we conclude that c is a concave transformation of $\hat{c}$ for HARA utility with $\gamma \geq -1$? In general, the answer is negative. After a small modification of the above lemma, it can be shown that there exists a concave function $\tilde{h}$ such that c and $\hat{c}$ are related as follows:

$$c(y) = \tilde{g}^{-1} \circ \tilde{h} \circ \tilde{g}\left(\hat{c}(y)\right),$$

[12] In fact, this statement is not only true for concavity and convexity, but more generally for any property defined with respect to the transformations g_λ and $\hat{g}_{\hat{\lambda},\hat{\xi}}$.

where $\tilde{g}(c) = \frac{1}{\mu}\left(\frac{\lambda q}{u'(c)} - 1\right)$ is increasing. If $\tilde{g}$ is an affine function (u is logarithmic utility), then one can easily verify that the composition $\tilde{g}^{-1} \circ \tilde{h} \circ \tilde{g}$ is concave whenever $\tilde{h}$ is concave. For the logarithmic case, c is hence a concave transformation of $\hat{c}$. In general, however, we cannot be sure that the composition $\tilde{g}^{-1} \circ \tilde{h} \circ \tilde{g}$ is concave when $\tilde{h}$ is concave.[13]

5 Quantitative analysis (preliminary)

This quantitative exercise serves two purposes. First, we extend our theoretical results for applied purposes. For example, recall that, the theoretical results compare two allocations that implement the same effort level. In a calibrated/estimated framework we show that the key result of complementarity of capital taxation and income tax progressivity extends to the case where effort is allowed to change between the two scenarios.

The second target of this exercise is to evaluate the effect of (optimal) capital taxation on labor income taxes *quantitatively*. In order to this, we use consumption and income data and postulate that the data is generated by the model with optimal capital taxation. In particular, we estimate some of the key parameters of the model by matching some joint moments of consumption and income in an appropriately cleaned cross-sectional data. Then, we use the estimated (and postulated) parameters and also solve the model with no capital taxes. The final outcome is a comparison of the optimal labor income taxes between the two scenarios.

5.1 Data

We use PSID (Panel Study of Income Dynamics) data for 1991 and 1992 as adapted by Blundell, Pistaferri and Preston (2008). This data file contains consumption data and income data at the household level. The consumption data is imputed using food consumption (measured at the PSID) and household characteristics using the CEX (Survey of Consumption Expenditure) as a basis for the imputation procedure. Household data is useful for two reasons: (i) Consumption can be credibly measured at the household level only. (ii) Taxation is mostly determined at the family level (which is typically equivalent to the household level) in the United States.

In our model, we have ex-ante identical individuals who face the same (partially endogenous) process of income shocks. In the data, however, income is influenced by factors such as age, education and innate ability. We want to control for these characteristics to obtain a clean measure

[13] Consider the following example: $\tilde{g}(c) = \exp(c)$ for $c > 0$, $\tilde{h}(x) = (x + d)^{\alpha}$ for $x > 0$, with $0 < \alpha < 1$, $d > 0$. Then $\tilde{h}(x)$ is concave in x, but $\tilde{g}^{-1} \circ \tilde{h} \circ \tilde{g}\,(c) = \alpha \log(\exp(c) + d)$ is convex in c.

of income. To do this, we postulate the following process for income:

$$y_t^i = \phi(X_t^i)\theta^i\varepsilon_\tau^i$$

where y_t^i is household i's income at time t, X_t^i are observable household characteristics (age, education and race of the household head), θ^i are unobservable fixed effects at the household level (like average ability of the couple) and ε_τ^i is a shock to the income of the household. Furthermore, we assume that this shock evolves according to a geometric random walk $\varepsilon_\tau^i = \varepsilon_{\tau-1}^i\eta_t^i$ (e.g., see Blundell et al. (2008)). The term η_t^i is our measure of the cleaned income shock.

In order to isolate η_t^i, we first run a regression of $\log(y_t^i)$ on X_t^i for $t = 1991$ and $t = 1992$. If we call the residual of this equation u_t^i, then this is our estimate of $\log(\theta^i\varepsilon_\tau^i)$. Now, to estimate the shock η_t^i, we use the equation $\varepsilon_\tau^i = \varepsilon_{\tau-1}^i\eta_t^i$ and set

$$\hat{\eta}_t^i = \exp\left(u_t^i - u_{t-1}^i\right).$$

The next objective is to find the consumption function. In order to relate it to the cleaned income measure η_t^i, we postulate that the consumption function is multiplicatively separable as well. Let $h_t^i := (\varepsilon_0^i, \varepsilon_1^i, ..., \varepsilon_t^i)$ be the history of individual income shocks up to period t:

$$\begin{aligned} c_t^i &= g^0(Z_t^i)g^1\left(\phi(X_t^i)\right)g^2\left(\theta^i\right)c\left(h_t^i\right) \\ &= g^0(Z_t^i)g^1\left(\phi(X_t^i)\right)g^2\left(\theta^i\right)c\left(h_{t-1}^i\right)c\left(\eta_t^i\right) \end{aligned}$$

where Z_t^i are household characteristics that affect consumption, but (by assumption) do not affect income, such as number of kids and household assets. Our target is to identify $c\left(\eta_t^i\right)$, the pure response of consumption to the income shock.

To isolate this effect, we first run a regression of $\log(c_t^i)$ on X_t^i and Z_t^i for $t = 1991$ and $t = 1992$. The residual of this equation is called v_t^i and estimates $\log\left(g^2\left(\theta^i\right)c\left(h_t^i\right)\right)$. Then, using the assumption $c\left(h_t^i\right) = c\left(h_{t-1}^i\right)c\left(\eta_t^i\right)$, we obtain our estimate of consumption in the data as:[14]

$$\hat{c}_t^i = \exp\left(v_t^i - v_{t-1}^i\right).$$

In the final stage, we use a flexible functional form to obtain $c(\cdot)$. In particular, we estimate the following regression:

$$\log(\hat{c}_t^i) = \sum_{j=0}^{4}\gamma_j\left(\log(\hat{\eta}_t^i)\right)^j.$$

[14]We are planning to use a more flexible form here: $\hat{c}_t^i = \exp\left(v_t^i - \beta v_{t-1}^i\right)$, where β will be estimated from regressions of c_t^i on c_{t-1}^i. This structure will be consistent with the history dependence generated by the dynamic moral hazard model and many other models of optimal contracting with observable assets.

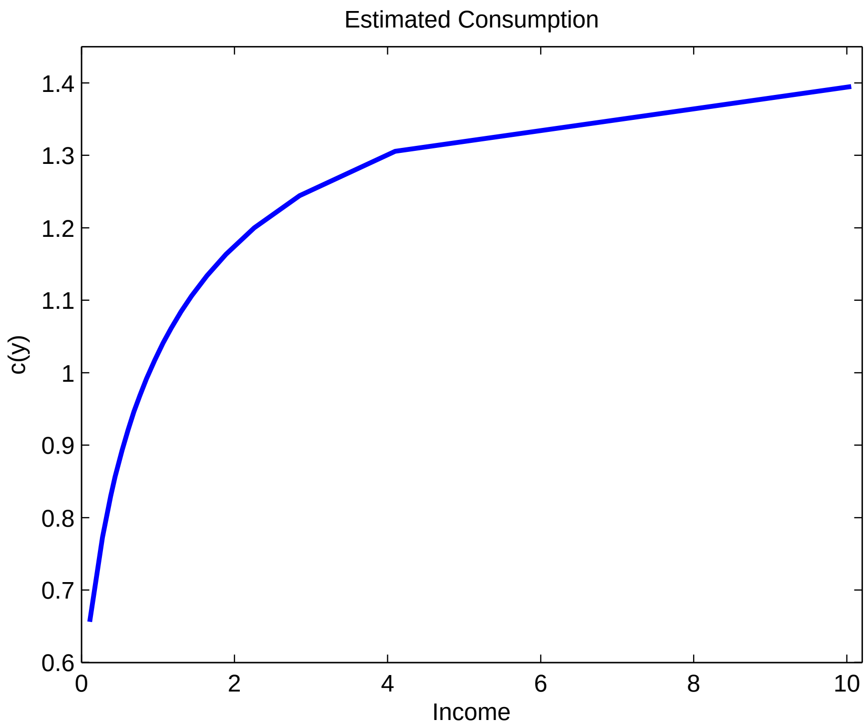

Figure 1: Estimate of consumption function from the data

Hence, in our model's notation, the estimate of the consumption function is given by

$$\hat{c}\,(y) = \exp\left(\sum_{j=0}^{4} \hat{\gamma}_j\,(\log(y))^j\right).$$

Figure 1 displays the estimated consumption function.

5.2 Estimation of model parameters

For the quantitative exploration of our model, we move to a formulation with discrete income levels. We assume that we have N levels of second-period income, denoted by y_i with $y_i > y_{i-1}$. This implies that the density function of income $f(x, e)$ is replaced by probability weights $p_i(e)$, with $\sum_{i=1}^{N} p_i(e) = 1$ for all e. For the estimation of the parameters, we impose further structure. We assume

$$p_i(e) = \exp(-\rho e)\pi_i^l + (1 - \exp(-\rho e))\pi_i^h,$$

where π_i^h and π_i^l are probability distributions. In addition to tractability, this formulation has the advantage that it satisfies the requirements for the applicability of first-order approach given by Abraham, Koehne and Pavoni (2010). Along these lines, we choose a quadratic cost function for effort, $v(e) = \alpha e^2$, and a CRRA utility function, $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$.

For now, we fix some parameters. In particular, we use $\alpha = 0.5$, $\beta = q = 0.96$ and $\sigma = 3$.[15] The remaining parameters of the model are ρ and G and the probability weights $\left\{\pi_i^h, \pi_i^l\right\}_{i=1}^N$ that determine the likelihood ratios. We use the optimality conditions to design a method of moments estimator for these parameters. In particular, we use $N = 20$ and choose the medians of the 20 percentile groups of cleaned income for the income levels $y_1, ..., y_{20}$. Our target moments are $p_i(e^*) = 1/20$ for all i, where e^* is the optimal effort, and $c_i^* = \hat{c}(y_i)$, where c_i^* is the optimal consumption in the model. We use the identity weighting matrix in the estimation.[16]

Since the probabilities π_i^l and π_i^h each sum up to one, we have $N-1$ parameters each. Moreover, we have to estimate the parameter ρ. To summarize, we have to estimate $2N-1$ parameters and use the following $2N-1$ model restrictions for these parameters:

$$
\begin{aligned}
p_i(e^*) &= \exp(-\rho e^*)\pi_i^l + (1-\exp(-\rho e^*))\pi_i^h \text{ for } i = 1, ..., N-1, && (16)\\
\frac{\lambda^*}{u'(c_i^*)} &= 1 + \mu^* \frac{\exp(-\rho e^*)\rho\left(\pi_i^h - \pi_i^l\right)}{p_i(e^*)} \text{ for } i = 1, ..., N. && (17)
\end{aligned}
$$

Recall that, whenever we impose optimal capital taxation, we have $\xi = 0$, which generates the simple form of the above equations. Notice that these equations also include e^*, μ^* and λ^*. For these variables, we use three optimality conditions, which we require to be satisfied exactly. First (17) implies that

$$\sum_{i=1}^N p_i(e^*)\frac{1}{u'(c_i^*)} = \frac{1}{\lambda^*}.$$

Then, we can use the first-order incentive compatibility constraint for effort,

$$2\alpha e^* = \beta\rho\exp(-\rho e^*)\sum_{i=1}^N \left(\pi_i^h - \pi_i^l\right) u(c_i^*), \tag{18}$$

together with the planner's first-order optimality condition for effort (I have used the fact that (18) cancels out):

$$\lambda^*\beta\rho\exp(-\rho e^*)\sum_{i=1}^N \left(\pi_i^h - \pi_i^l\right)(y_i - c_i^*) = \mu^*\left(2\alpha + \beta\rho^2\exp(-\rho e^*)\sum_{i=1}^N \left(\pi_i^h - \pi_i^l\right) u(c_i^*)\right).$$

[15] We have made some sensitivity analysis with respect to the risk aversion parameter. Our results are qualitatively the same for the range $\sigma \in [1, 5]$, but, as we will see, the differences between the two scenarios are more pronounced if risk aversion is bigger.

[16] This choice turned out to be irrelevant, because we obtained a practically perfect fit.

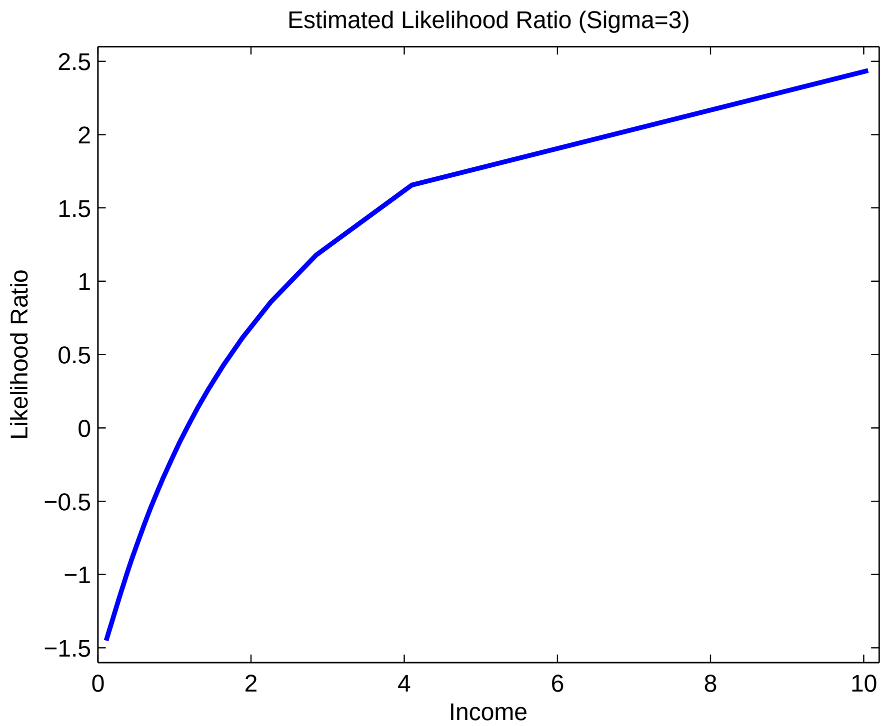

Figure 2: Estimated Likelihood Ratio

Then, normalizing $b_0 = 0$, we obtain from the government's budget constraint the implied government consumption as

$$G = y_0 - c_0^* + q \sum_{i=1}^{N} p_i(e^*)(y_i - c_i^*). \tag{19}$$

To solve (19) for G, we assume $y_0 = \sum_{i=1}^{N} p_i(e^*) y_i$ and use the first-order condition with respect to c_0,

$$\frac{1}{u'(c_0^*)} = \frac{1}{\lambda^*}.$$

The estimated ρ parameter is 9.52, while G is given by 1.2236. This represents around 37.5% of government consumption as of per period output. We plot the the estimated likelihood ratio on Figure 2. The figure brings good news, since given our functional forms a monotone likelihood ratio guarantees the validity of the FOA. The likelihood ratio is concave as well. This implies that the first part of Proposition 3 is applicable in the optimal capital tax case, hence consumption has to be concave, which is indeed the case.

5.3 Results

We use the preset and estimated (calibrated) parameters of the model above to determine the optimal allocation for the scenario where no capital taxes are allowed. Figure 3 displays the optimal second-period consumption allocation for this scenario.

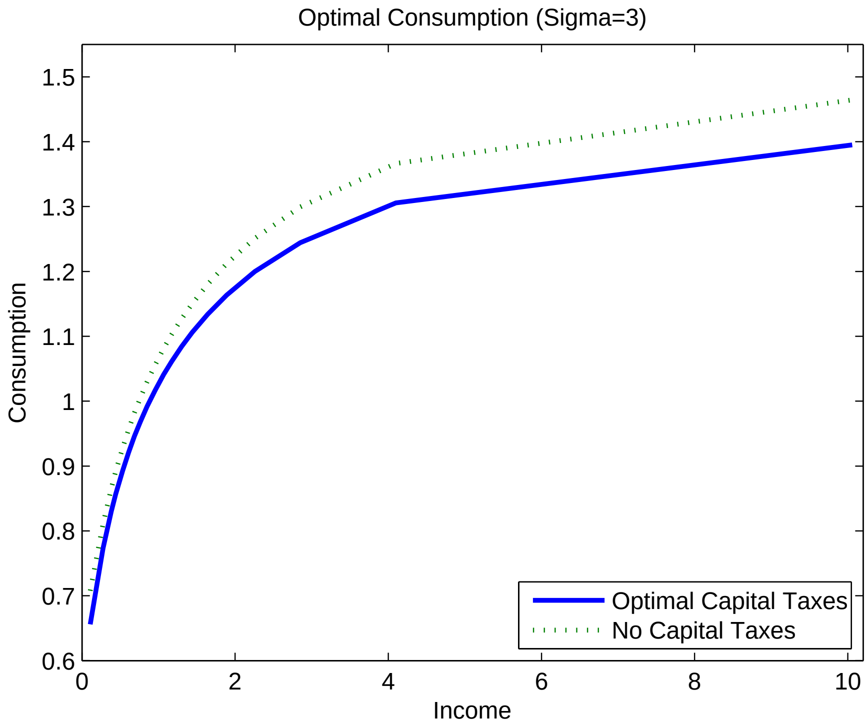

Figure 3: Optimal consumption with and without capital taxation

It is obvious from the picture that the average level of second-period consumption is higher in the case without capital taxation. The fact that consumption should be more frontloaded in the case with capital taxation is a well known result and very intuitive given that the optimal capital tax is always positive.

We also observe that - since consumption is concave for the two cases - optimal labor income taxes are progressive in both scenarios. Recall that the discrete version of our marginal tax takes the form $1 - (c_{i+1} - c_i)/(y_{i+1} - y_i))$. First note that since the estimated likelihood ratio is concave we can invoke the first part of the Corollary to Proposition 6 which states that if $\hat{c}$ is g_λ-concave then c is g_λ-concave. Moreover, for $\sigma = 3$, $g_\lambda = \lambda c^3$ is convex, hence g_λ-concavity implies concavity. However, recall that, for these computations we did not fix effort to be the same across the two allocations which was a requirement for Proposition 6. On the one hand, this results show that the endogenous response of effort to the lack of capital taxation does not affect the qualitative results (at least for this set of parameters). On the other hand, we will also show below that the changes in effort (and consequently the likelihood ratio) has a non-negligible quantitative effect.

In order to be able to compare progressivity across the two scenarios quantitatively, we have created two measures. The first one simply compares the marginal tax rate at the top level of income with that of the bottom level of income. The other one measures the average slope of the marginal tax function across the two scenarios where the weights are the population weights of the different income levels. A higher average slope obviously indicates a higher degree of progressivity. The results are displayed in Table 1. The results show a clear pattern. Progressivity is reduced when no capital taxes are imposed. However, this differences are negligible for log utility, significant but modest for our benchmark case and very large (progressivity is halved) for $\sigma = 5$. Note that these values are all in the empirically plausible range of risk aversion parameters. Hence, the quantitative significance of the lack of capital taxation still requires some further research. In particular, an extension of our empirical procedure is needed to be able to pin down σ. We obtain a similar message if we consider the welfare losses due to not taxing capital in consumption equivalent terms (presented in the last row of Table 1). The losses are negligible for the log case, but considerable for higher values of risk aversion.

Table 1: Quantitative Measures of Progressivity and Welfare Losses

σ	1	2	3	4	5
	Average Slope of Marginal Tax				
Optimal Capital Taxes	0.322	0.321	0.321	0.320	0.321
No capital Taxes	0.318	0.303	0.280	0.245	0.157
	Difference between Top and Bottom Bracket				
Optimal Capital Taxes	0.684	0.684	0.684	0.684	0.681
No capital Taxes	0.677	0.647	0.597	0.524	0.348
	Welfare Losses from Not Taxing Capital (%)				
	0.01	0.10	0.34	0.79	1.20

We can get some intuition why the differences are increasing in σ by examining equation (12) for our case:

$$\hat{\lambda}\hat{c}_i^{\sigma} - \hat{\xi}\frac{\sigma}{\hat{c}_i} = 1 + \hat{\mu}\frac{\exp(-\rho\hat{e})\rho\left(\pi_i^h - \pi_i^l\right)}{p_i(\hat{e})} \qquad \text{for } i = 1, ..., N,$$

The direct effect of no capital taxation is driven by $\hat{\xi}a(\hat{c})$. Note that the higher is σ, the higher is the discrepancy between the Euler equation characterizing the no capital taxation case and the inverse Euler characterizing the optimal capital taxation case. This will imply that $\hat{\xi}$ is increasing with σ. Moreover, absolute risk aversion is given by $\sigma/\hat{c}$, which is also increasing in σ.

Finally, we would like to relate the quantitative results to Proposition 5 as well. There we have shown that under convex absolute risk aversion, whenever consumption is concave function of the likelihood ratio in the no capital tax case, the same must hold in the model with optimal capital taxes. Recall that this result was obtained assuming constant effort levels across the two scenarios. Therefore we compute the optimal allocation for the scenario without capital taxation given the effort level from the optimal capital tax case. Intuitively, we disregard the planner's optimality condition regarding effort. Figure 4 displays the results of these calculations as a function of the likelihood ratio, which is (by construction) the optimal likelihood ratio under optimal capital taxes.

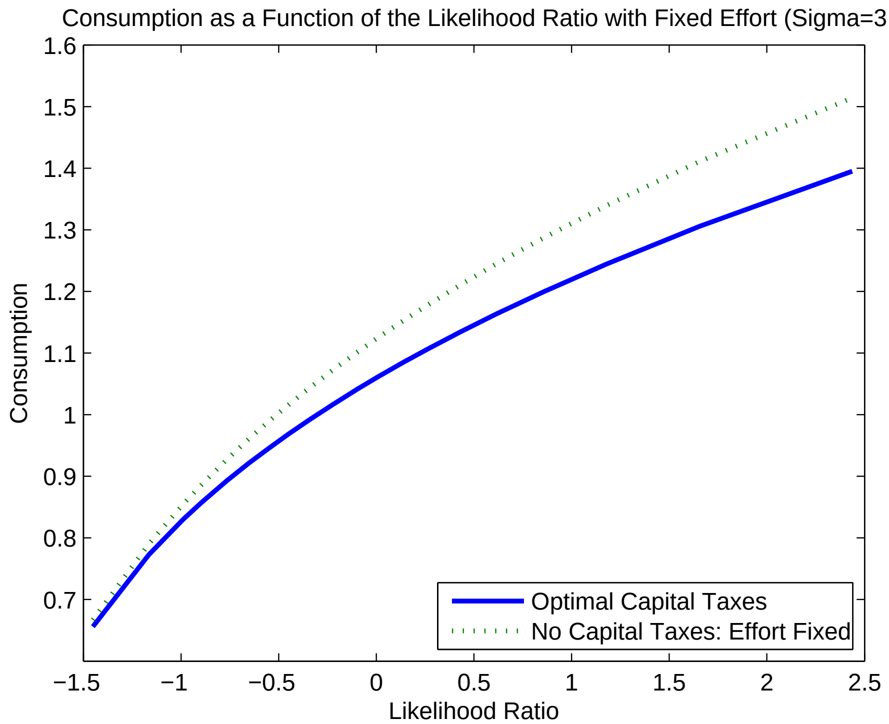

Figure 4: Optimal Consumption as Function of Likelihood Ratio (Fixed Effort)

This figure is clearly in line with the theoretical results of Proposition 5, as the consumption allocation under optimal capital taxation is a more concave function of the likelihood ratio (this eyeball check holds true when we use can be shown .

6 Conclusions and Outlook

This paper analyzed how capital taxation changes the optimal tax code on labor income. Whenever preferences exhibit convex absolute risk aversion, we found that optimal consumption moves in a 'more concave' way with labor income when capital is taxed. In this sense, labor income taxes become more progressive when capital is taxed. We complemented our theoretical results with a quantitative analysis based on individual level US consumption and income data.

The model we presented here is one of action moral hazard. This is mainly done for tractability. Although a more common interpretation of this model is that of insurance, we believe that it conveys a number of general principles for optimal taxation that are also valid in models of ex-ante redistribution. Of course, the quantitative analysis might change a lot in this case. A recent example of a quantitative exploration of optimal income taxation under observable assets (in two periods) is Golosov, Troshkin, and Tsyvinski (2009). They, however, do not study how capital taxation affects the optimal tax scheme in presence of asset accumulation. This is hence a possible direction for further research that can be addressed by including ex-ante (unobservable) heterogeneity into our model.[17]

Some further extensions of the model could be useful for future research. For example, the study of assets different from the bond might highlight the boundaries of the present findings.

References

[1] Abraham, A., S. Koehne, and N. Pavoni (2010): "On the First-Order Approach in Principal-Agent Models with Hidden Borrowing and Lending," *mimeo*, University College London and European University Institute.

[2] Albanesi, S., and C. Sleet (2006): "Dynamic Optimal Taxation with Private Information," *Review of Economic Studies*, 73(1), 1–30.

[3] Avriel, M., W. E. Diewert, S. Schaible, and I. Zang. (1988): *Generalized Concavity*. Plenum Publishing Corporation, New York.

[4] Blundell R, Pistaferri L., Preston I. (2008) 'Consumption inequality and partial insurance.' *American Economic Revew*, 98: 1887–921.

[17] As mentioned in footnote 2, the case with observable heterogeneity can be handled quite easily and none of our results would change.

[5] Conesa, J. C., and D. Krueger (2006): "On the optimal progressivity of the income tax code," *Journal of Monetary Economics*, 53(7), 1425–1450.

[6] Conesa, J. C., S. Kitao, and D. Krueger (2009): "Taxing Capital? Not a Bad Idea After All!," *American Economic Review*, Vol. 99 (1), 25-48.

[7] Domeij, D., and M. Floden (2006): "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased," *Review of Economic Dynamics*, 9(2), 242–262.

[8] Eaton, J., and H. S. Rosen (1980): "Optimal Redistributive Taxation and Uncertainty," *The Quarterly Journal of Economics*, 95(2), 357–64.

[9] Gollier, C. (2001): *The Economics of Risk and Time*. MIT Press.

[10] Golosov, M., and A. Tsyvinski (2006): "Designing Optimal Disability Insurance: A Case for Asset Testing," *Journal of Political Economy*, 114(2), 257–279.

[11] Golosov, M., M. Troshkin, and A. Tsyvinski (2009), "Optimal Dynamic Taxes," *mimeo*.

[12] Gottardi, P., and N. Pavoni (2010): "Ramsey Asset Taxation under Asymmetric Information," *mimeo,* European University Institute.

[13] Guiso, L., and M. Paiella (2008): "Risk Aversion, Wealth, and Background Risk," *Journal of the European Economic Association*, 6(6), 1109–1150.

[14] Guvenen, F., B. Kuruscu, and S. Ozkan (2009): "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," *NBER Working Papers 15526*, National Bureau of Economic Research, Inc.

[15] Holmstrom, B. (1979): "Moral hazard and observability," *Bell Journal of Economics*, 10(1), 74–91.

[16] Kocherlakota, N. R. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73(5), 1587–1621.

[17] Laroque, G. (2010): "On Income and Capital Taxation in a Life Cycle Model with Extensive Labor Supply," *mimeo* University College London.

[18] Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(114), 175–208.

[19] Sadka, E. (1976): "On Progressive Income Taxation," *American Economic Review*, 66(5), 931–35.

[20] Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), 205–29.

[21] Varian, H. R. (1980): "Redistributive taxation as social insurance," *Journal of Public Economics*, 14(1), 49–68.