Working Paper Series

Optimal Wealth Taxes with Risky Human Capital

WP 05-13

Borys Grochulski
Federal Reserve Bank of Richmond

Tomasz Piskorski
New York University

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Optimal Wealth Taxes with Risky Human Capital*

Borys Grochulski
Federal Reserve Bank of Richmond
borys.grochulski@rich.frb.org

Tomasz Piskorski
New York University
tp431@nyu.edu

Federal Reserve Bank of Richmond Working Paper No. 05-13
December 2005

Abstract

We study the structure of optimal wealth and labor income taxes in a Mirrlees economy in which the productivity of labor (i.e., skill) is private, stochastic, and endogenous. Individual agents' skills are determined by their level of human capital. Human capital is not publicly observable and the returns to human capital investment are subject to idiosyncratic shocks. Preferences are not assumed to be additively separable in consumption and human capital investment and, thus, the intertemporal marginal rates of substitution of consumption are private information. We characterize the optimal allocation and a tax system that implements this allocation in equilibrium. The optimal allocation does not satisfy the "reciprocal Euler equation" of Rogerson [Econometrica, 1985], which holds in Mirrlees economies with exogenous skills. The tax system we use in our decentralization of the optimum consists of a wealth tax that is linear in wealth and a labor income tax that depends solely on labor income. The result of Kocherlakota [Econometrica, 2005], establishing the optimality of zero expected marginal wealth tax rate, holds in our model. We show that endogenous skill determination affects the volatility of marginal wealth taxes rather than their expectation. Relative to economies with exogenous skills, the optimal marginal wealth tax rate is more volatile in our endogenous skill economy. Also, we demonstrate the optimality of a wedge in the returns on the two assets present in our economy: At the optimum, the marginal return on human capital investment is strictly larger than the marginal return on physical capital investment.

Keywords: Optimal taxation, human capital, Mirrlees approach.
JEL classification: E62, H21, J24.

1 Introduction

Recent literature obtains important results characterizing optimal capital and labor income taxes in dynamic Mirrlees economies.[1] In a Mirrlees economy, agents are affected by idiosyncractic, privately

*We would like to thank Stefania Albanesi, David Backus, Alberto Bisin, Gian Luca Clementi, Douglas Gale, Mikhail Golosov, Lars Peter Hansen, Narayana Kocherlakota, Christopher Phelan, Edward Simpson Prescott, Thomas Sargent and seminar participants at NYU, the Federal Reserve Bank of Richmond, the 2005 SAET conference, and the 2005 Chicago-NYU Workshop for helpful comments and suggestions. All remaining errors are ours. The views expressed here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] See Kocherlakota (2005a) for a review.

observable shocks to the productivity of their labor effort. In the Mirrlees approach to optimal taxation, the role of the tax system is to fund government purchases and insure the productivity risk. The optimal taxation problem is to characterize a tax system that fulfills this dual role efficiently, given the informational constraints imposed by the lack of public observability of the idiosyncractic productivity shocks.

In a ground-breaking paper, Mirrlees (1971) solves the optimal taxation problem in a static setting. Taking all but the labor effort decisions as given, he characterizes optimal labor income taxes. The main limitation of the static approach is that it ignores the effect that taxes have on agents' investment decisions. The contribution of the recent literature is in the characterization of optimal tax systems in dynamic settings in which agents' physical capital investment decisions (i.e., savings) are endogenous.

Physical capital investment, however, is not the only important category of investment decision problems that agents face over the life cycle. There is ample evidence suggesting that human capital investment decisions are at least equally important.[2] By investing in their human capital, people affect profoundly their future skills, and, consequently, their wages, earnings and welfare. The existing literature on optimal taxation with endogenous savings, however, takes the evolution of agents' skills as exogenous and thus ignores the effect that taxes have on agents' human capital investment decisions. In this paper, we solve the optimal taxation problem in a simple dynamic setting in which both the physical capital and human capital investment decisions are endogenous.

Introducing endogenous human capital into a two-period Mirrlees economy, we make two assumptions about the technology of human capital formation. First, we assume that agents' productivity in the second period depends on the amount of resources they devote to human capital accumulation in the first period. This assumption is standard in the human capital literature [see, e.g., Heckman (1976), Boldrin and Montes (2005)].[3] Second, we assume that the return on human capital investment is subject to stochastic idiosyncractic shocks. This assumption is well documented in empirical studies [see, e.g., Palacios-Huerta (2003)].

The distribution of information in our model is as follows. Following the optimal taxation literature surveyed in Kocherlakota (2005a), we take physical capital, output, and labor income as observable. As a basic feature of all Mirrlees economies, individual skills are agents' private information. Therefore, we assume that agents' human capital is not publicly observable. Finally, we need to make an assumption about the observability of human capital investment. We assume that it is not publicly observable how much resources an agent spends on human capital investment and how much on consumption and, therefore, both are private information of the agent.

The difficulty in distinguishing between human capital investment and consumption expenditures has been long recognized by academic economists.[4] It has also been recognized as a problem in the ongoing policy debate on how to design the tax system in order to foster human capital accumulation.[5] At the core of this measurement problem lies the fact that, in reality, there is a human

---

[2] Some estimates put the value of human capital at 93% of all wealth in the US. See Palacios-Huerta (2003a) and references therein.

[3] Many microeconomic studies also assume that effort is an input in the technology of human capital production. We do not include this input in our production function for clarity of exposition. Our results do not depend on this abstraction.

[4] Theodore Schultz in his 1961 Presidential Address to the AEA stresses that "Much of what we call consumption constitutes investment in human capital." See Schultz, (1961, 1961a) and Shaffer (1961) for an extensive discussion, and Davies, Zeng and Zhang (2000) for a recent application.

[5] A 2005 memorandum to the President's Advisory Panel on Federal Tax Reform on tax treatment of investment

capital investment dimension to ordinary consumption and a significant amount of consumption value in human capital investment activities such as education and training. Agents use a large variety of goods, services and non-market activities as vehicles for their human capital investment and consumption. It is difficult to measure the relative "loadings" of human capital investment and pure consumption embedded in a particular good or service. In a model with a single consumption good, the assumption that consumption and human capital investment are indistinguishable to an outside observer captures this measurement problem better than the alternative assumption of full observability.

Building on these assumptions, we fully specify an economic environment with endogenous, risky, and private human capital; characterize the optimal allocation of resources; and construct a tax system that implements this allocation in equilibrium.

The incentive problem that arises in our environment is considerably different from the incentive problems that shape optimal capital and labor income taxes in dynamic Mirrlees economies with exogenous skills. In these economies [see Albanesi and Sleet (2005) and Kocherlakota (2005)], taxes, in addition to raising revenue, must provide incentives to prevent highly productive agents from pretending to be low-skilled, i.e., from shirking. In our model, agents can end up highly productive ex post only if they make sufficient human capital investment ex ante. Taxes, therefore, must provide incentives not only to ensure that agents do not shirk ex post, but also to induce the agents to make efficient investment in their human capital ex ante.

We demonstrate how this incentive problem can be overcome by a capital and labor income tax system similar to that of Kocherlakota (2005), in which capital taxes are linear in capital and labor income taxes depend only on labor income. As in Albanesi and Sleet (2005) and Kocherlakota (2005), the optimal marginal capital tax rate is uncertain ex ante, i.e., at the time of investment. Ex post, the marginal capital tax rate is positive for agents with low labor income and negative for agents whose labor income is high. As in Kocherlakota (2005), the average optimal marginal capital tax rate is zero, which implies that the government collects no revenue from taxation of capital.

Given that we use the same tax system as Kocherlakota (2005), these results would not be surprising at all if it were not for one quite fundamental difference between our environment and the environment of Kocherlakota (2005). In our environment, agents' intertemporal marginal rates of substitution (IMRS) are agents' private information. As a consequence, the optimal allocation does not satisfy the so-called reciprocal Euler condition of Rogerson (1985) and Golosov, Kocherlakota and Tsyvinski (2003), which equates the inverse of the marginal utility of consumption at date t to the discounted expected value of the inverse of the marginal utility of consumption at date $t+1$. In the environment of Kocherlakota (2005), despite private productivity shocks, the IMRS are publicly known and the optimal allocation does satisfy a generalized version of the Rogerson condition. The proofs of the decentralization result and the zero average capital tax result given in Kocherlakota (2005) rely on this property of the optimal allocation. In our model, we obtain the decentralization and zero average tax results despite the fact that the IMRS are private and the Rogerson condition does not hold.

What, then, does our model say about the impact of endogenous human capital on the structure of optimal capital taxes? Endogenous human capital does not affect the average optimal tax rate,

---

in human capital prepared by the Treasury Department's Office of Tax Analysis says, "In practice, it can be very difficult to distinguish between human capital investment and education consumption." See the reference United States Department of Treasury, Office of Tax Analysis (2005) for a full discussion.

which continues to be equal to zero. Instead, it turns out in our model that allowing for endogenous determination of skills increases the volatility of optimal marginal tax rates.

This result follows from the structure of the incentive problem that shapes optimal taxes in our model. Allowing for endogenous human capital accumulation through unobservable investment decisions adds an extra dimension to the space of strategies that agents can use to deviate from the socially optimal pattern of investment, labor and consumption. With such an enhanced set of deviation opportunities, the incentive problem of our environment is more severe, relative to environments in which skills are exogenous. This translates, at the optimum, into a larger intertemporal wedge between the shadow interest rate of consumption and the rate of return on physical capital investment. In order to support this wedge in equilibrium, capital taxes have to introduce more risk into the return on physical capital investment, i.e., the ex post marginal capital tax rates have to be more volatile.

The presence of two assets that can be used as a vehicle for an intertemporal transfer of resources is another important difference between our environment and the exogenous-skill environments studied in the literature. In our model, in addition to physical capital, human capital can be used indirectly to transfer resources across time. Comparing the optimal rates of return on these two assets, we identify a wedge in asset return rates that implies the optimality of a human capital premium. Namely, it turns out to be optimal for the return on human capital investment to be strictly larger than the return on physical capital investment. The reason for the optimality of this wedge is the difference between the social costs of human and physical capital investments. As human capital is unobservable, human capital investment bears a larger incentive cost than the observable physical capital investment. At the optimum, this larger social cost has to be offset by a larger return on human capital investment.

As the above discussion makes apparent, Kocherlakota (2005) is the paper in the literature that is the most closely related to ours. Another paper that uses the tax system of Kocherlakota (2005) is Farhi and Werning (2005). This paper studies optimal estate taxation in an environment in which the consumption IMRS are public. The optimal allocation studied in Farhi and Werning (2005) does not satisfy the Rogerson condition and the average optimal capital (i.e., estate) tax rate is not zero. Also, Albanesi (2005) shows in a model with entrepreneurial capital and moral hazard in which the optimal allocation does satisfy the Rogerson condition that the average optimal marginal tax rates on all assets are zero. The result of Farhi and Werning (2005) together with Albanesi (2005) and the original result of Kocherlakota (2005) suggest that, in a class of tax systems that are linear in capital, the zero average capital tax result holds if and only if the Rogerson condition holds at the optimal allocation. The results of our paper contradict this intuition. We do not have the Rogerson condition but capital taxes are zero on average. Albanesi and Sleet (2005) show that capital taxes may be non-zero even in environments in which the Rogerson condition holds at the optimum. Their tax system, however, is in general not linear in wealth and, thus, their results concern a different decentralization than the one used in Kocherlakota (2005), Farhi and Werning (2005), Albanesi (2005) and our paper.

Kocherlakota (2004) gives an example of an environment in which the consumption IMRS are private and optimal capital taxes have to be nonlinear in capital, i.e., a decentralization with linear capital taxes does not exist. In this example, preferences are not additively separable in consumption and labor. In our model, the IMRS are private but a decentralization with linear capital taxes exists. This is because the preferences we assume are non-separable in consumption and human capital investment but remain separable in consumption and labor.

The question of optimal taxation in a model with human capital accumulation has been addressed in many papers in the context of the so-called Ramsey approach to optimal taxation (see for example Jones, Manuelli and Rossi (1997)). In this approach, the government is restricted to use linear taxes and this restriction is not derived from the fundamentals of the model. Our paper is different, as we use the Mirrlees approach. We specify explicitly an informational friction in the fundamentals of the model that endogenously constrains the set of tax instruments that the government can use. Kapicka (2005) is a related paper that studies taxation of human capital in a Mirrlees environment. This paper, however, imposes an exogenous restriction on the set of tax instruments that the government can use, and thus derives a tax system that is constrained-suboptimal. In addition, Kapicka (2005) disallows capital markets. In our paper, we allow both human and physical capital accumulation and derive a tax system that is optimal, relative to the informational frictions of our environment.

This paper is organized as follows. Section 2 defines the environment. Section 3 provides a characterization of the optimal allocation. Section 4 provides a decentralization theorem and proves the zero average tax result. Section 5 shows the effect of endogenous human capital on the volatility of optimal capital taxes. Section 6 provides some numerical examples. Section 7 concludes.

2 Environment

Consider a two-date ($t = 0, 1$) economy populated by a continuum of ex ante identical agents. The size of the population is normalized to unity. There is a single consumption good at each date. In period 0 each agent is endowed with $k_0 > 0$ units of capital, which can be invested or consumed.

Preferences: Agents' preferences over stochastic streams of consumption $c = (c_0, c_1)$ and labor effort at $t = 1$, l, are given by

$$u(c_0) + \beta E\left[u(c_1) + v(l)\right], \tag{1}$$

where $u : \mathbb{R}_+ \to \mathbb{R}$ is a strictly increasing, strictly concave C^2 function and $v : \mathbb{R}_+ \to \mathbb{R}$ is a strictly decreasing, strictly concave C^2 function such that $v(0) = 0$.

Technology: Using an available 1 to 1 technology, agents can transform the date-zero capital into consumption c_0, date-one physical capital k_1, or into a human capital investment good i. The human capital production technology is stochastic. A date-zero human capital investment of size $i \geq 0$ produces the amount h of date-one human capital given by

$$h = \theta i,$$

where $\theta \in \Theta \equiv \{0, 1\}$ is an i.i.d. across agents shock to human capital investment productivity. The probability of $\theta \in \Theta$ is denoted by $\pi(\theta)$, with $0 < \pi(\theta) < 1$ for both $\theta \in \Theta$. Individual realizations of this shock are private. We assume that the exact Law of Large Numbers applies, so $\pi(\theta)$ also represents the fraction of agents whose shock realization is θ.

Human capital determines agents' productivity at date one. An agent whose human capital level is h and who works l units time provides y units of effective labor. The amount of y, as a function of h and l, is given by

$$y = f(h)l,$$

where $f : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing, strictly concave, differentiable, publicly known function satisfying the Inada conditions: $\lim_{h\to 0} f'(h) = +\infty$, $\lim_{h\to\infty} f'(h) = 0$.

Effective labor y is publicly observable. However, the individual amount invested i, stock of human capital h, and labor l are private information of each agent.

In period 1, the consumption good is produced according to an aggregate production function $F : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$, which has aggregate physical capital K and aggregate effective labor Y as inputs. The amount of the consumption good produced at $t = 1$ from inputs K and Y is given by

$$F(K,Y) = Z(K,Y) + (1-\delta)K.$$

We assume that Z is strictly increasing, strictly concave, publicly known, constant returns to scale, and C^2. Also, we assume that Z satisfies the following Inada conditions:

$$\lim_{K\to 0} Z_1(K,Y) = \infty, \;\; \lim_{Y\to 0} Z_2(K,Y) = \infty,$$
$$Z(K,0) = Z(0,Y) = 0,$$

where Z_i denotes the first partial derivative of Z with respect to the i-th argument. The parameter $\delta \in [0,1]$ is the depreciation rate of physical capital. Undepreciated capital is available for consumption at date $t = 1$.

Information: Publicly observable are each agent's capital holdings k_0, k_1 and effective labor y. Since there is no investment in human capital at time $t = 1$, period-1 consumption c_1 is also publicly known. Human capital investment i, consumption at time zero c_0, the stock of human capital h, and the labor effort at time $t = 1$, l, are private.

Definition 1 *A (type-identical) allocation in this economy is a collection* (c, i, k_1, l, K, Y) *where*
$c = (c_0, c_1)$*, with* $c_0 \in R_+$ *and* $c_1 : \Theta \to R_+$*, denotes consumption,*
$i \in R_+$ *denotes investment,*
$k_1 \in R_+$ *is the physical capital held by an individual agent at* $t = 1$*,*
$l : \Theta \to R_+$*, is an individual agent's labor,*
$K \in R_+$ *is the aggregate capital stock at* $t = 1$*,*
$Y \in R_+$ *is the aggregate effective labor input at* $t = 1$*.*

Definition 2 *An allocation* (c, i, k_1, l, K, Y) *is resource feasible (RF) if*

$$c_0 + i + k_1 \leq k_0, \tag{2}$$
$$\sum_{\theta\in\Theta} \pi(\theta) c_1(\theta) + G \leq F(K,Y), \tag{3}$$
$$K = k_1, \tag{4}$$
$$Y = \sum_{\theta\in\Theta} \pi(\theta) f(\theta i)\, l(\theta), \tag{5}$$
$$c_0, c_1(\theta), i, k_1, l(\theta) \geq 0 \textit{for all } \theta \in \Theta\textit{, with } k_0 > 0, G \geq 0 \textit{given.}$$

The condition (2) is the resource constraint of time $t = 0$ and the condition (3) is the resource constraint of time $t = 1$. G is government spending at time $t = 1$. Equations (4) and (5) are accounting identities.

Definition 3 *An allocation* (c, i, k_1, l, K, Y) *is incentive compatible (IC) if*

$$u(c_0) + \beta \sum_{\theta \in \Theta} \pi(\theta) \left[u(c_1(\theta)) + v(l(\theta)) \right] \geq V(c, i, l), \tag{6}$$

where

$$V(c, i, l) = \max_{i'} u(c_0 + i - i') + \beta \sum_{\theta \in \Theta} \pi(\theta) w(\theta, i') \tag{7}$$

$$s.t.\ 0 \leq i' \leq c_0 + i \tag{8}$$

with

$$w(0, i') = \max \left\{ u(c_1(0)) + v(l(0)),\ u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(0)} \right) \right\}, \tag{9}$$

$$w(1, i') = \max \left\{ u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i')} \right),\ u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i')} \right) \right\}. \tag{10}$$

Formally, an incentive compatible allocation is an outcome of an incentive compatible direct revelation mechanism. The interpretation is as follows. At $t = 0$, a social planner makes a recommendation about the choices of c_0, i, and k_1. At $t = 1$, agents observe their individual realizations of the human capital investment shock θ and reveal them to the planner. As a function of those revealed realizations, the planner makes effective labor and consumption assignments. Agents provide labor required and consume in amounts assigned by the planner. The presence of private information is a restriction on the set of allocations that can be chosen by the planner. Agents must be willing to (a) follow the planner's human capital investment recommendation at $t = 0$, and (b) truthfully reveal θ at $t = 1$. The function V in (7) gives the maximal utility that agents can attain by deviating from the planner's investment recommendation and/or lying about their realizations of θ. The IC constraint (6) requires the planner to use such allocations that agents do not have an incentive to deviate from the planner's recommendation or to lie about θ. By the Revelation Principle, the restriction of the set of allocations to the set outcomes of direct revelation mechanisms is without loss of generality.

The formula (9) is the ex post continuation value $w(\theta, i')$ of an agent who invested i' in human capital and whose $\theta = 0$. This formula reflects the fact that the maximum continuation utility that can be achieved at $t = 1$ by an agent whose $\theta = 0$ is the larger of the two utility values attained by the two ex post reporting/labor strategies available to the agent. One is to truthfully report $\theta = 0$ and work $l(0)$ hours, which yields utility $u(c_1(0)) + v(l(0))$. The other is to pretend that $\theta = 1$ and work $\frac{f(i)l(1)}{f(0)}$ hours to provide the effective labor of the high-skilled equal to $f(i)l(1)$. This latter strategy yields utility $u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(0)}\right)$. Note that when $f(0) = 0$ and $f(i)l(1) > 0$, agents whose $\theta = 0$ cannot pretend to be high-skilled, as their skill level is zero. In this special case, therefore, $l(0) = 0$ at any efficient allocation, and (9) reduces to $w(0, i') = u(c_1(0))$.

The expression (10) represents the ex post continuation value $w(\theta, i')$ of an agent who invested i' in human capital and whose $\theta = 1$. Similar to the case of $\theta = 0$, this continuation value is equal to the larger of the utility values attained by the two reporting/labor strategies available to the agent ex post. One is to pretend that $\theta = 0$ and work $\frac{f(0)l(0)}{f(i')}$ hours to provide the effective labor of the low-skilled $f(0)l(0)$, which yields utility $u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i')}\right)$. The other strategy is to truthfully report $\theta = 1$ and work $\frac{f(i)l(1)}{f(i')}$ hours to provide the effective labor of the high-skilled, which leads to utility $u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i')}\right)$. Note that under either strategy the disutility of labor of an agent whose $\theta = 1$ depends on the amount i' invested at $t = 0$. Because this amount is private, human capital investment is a hidden state variable and thus there is no common knowledge of preferences over continuation contracts in our model.

Definition 4 *An allocation (c, i, k_1, l, K, Y) is incentive feasible if it is incentive compatible and feasible.*

Definition 5 *An allocation (c, i, k, l, K, Y) is constrained optimal if it is incentive feasible and if it maximizes, in the class of all incentive feasible allocations, the ex ante expected utility of the representative agent.*

By Definition 5, an allocation is constrained optimal, if it is a solution to the social planner's problem defined as follows:

Problem P1

$$\max_{(c,i,k_1,l,K,Y)} u(c_0) + \beta \sum_{\theta \in \Theta} \pi(\theta) \left[u(c_1(\theta)) + v(l(\theta))\right]$$
$$s.t.\ (RF),\ (IC).$$

Due to strict concavity of preferences, there is a unique constrained optimal allocation in this economy. Due to the assumed Inada conditions, the optimum is interior. We denote the constrained optimal allocation by $(c^*, k_1^*, i^*, l^*, K^*, Y^*)$ and refer to it simply as the optimal allocation, or the optimum.

3 Characterization of the optimal allocation

In this section, we provide a characterization of the optimal allocation. In the first subsection, by identifying and disregarding the non-binding IC constraints, we reduce the complicated IC constraints of Definition 3 to a single equality constraint. We proceed in two steps. First, in Lemma 1, we simplify the IC conditions of Definition 3 without relaxing them. This allows us to rewrite the social planner's problem P1 in an equivalent form in which the IC conditions are reduced to two inequalities and one equality constraint. Then, in Lemma 2, we show that only one of those constraints binds at the optimal allocation. This constraint turns out to be the IC constraint that restricts the set of allocations available to the planner to such allocations at which agents cannot benefit from a joint deviation of under-investing in human capital ex ante and shirking ex post.

In the second subsection, we turn to the properties of the optimal allocation. The most significant feature of the optimal allocation is the fact that the so-called reciprocal intertemporal first order condition of Rogerson (1985), Golosov, Kocherlakota and Tsyvinski (2003), and Kocherlakota (2005) does not hold at the optimum. Instead, Propositions 1 and 2 provide two "modified Rogerson" conditions that characterize our optimal allocation. As a corollary, we obtain that it is optimal for the marginal return on human capital investment to exceed the marginal return on physical capital investment.

3.1 Simplifying the IC constrains

Lemma 1 *An allocation* (c, i, k_1, l, K, Y) *is incentive compatible if and only if*

$$u(c_1(0)) + v(l(0)) \geq u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(0)}\right), \tag{11}$$

$$u'(c_0) = -\pi(1)\beta v'\left(l(1)\right)l(1)\frac{f'(i)}{f(i)}, \tag{12}$$

and

$$u(c_0) + \beta\pi(1)\left[u(c_1(1)) + v(l(1))\right] \geq$$
$$u(c_0 + i - \widetilde{i}) + \beta\pi(1)\left[u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(\widetilde{i})}\right)\right], \tag{13}$$

where $\widetilde{i}$ *solves the following equation*

$$u'(c_0 + i - \widetilde{i}) = -\pi(1)\beta v'\left(\frac{f(0)l(0)}{f(\widetilde{i})}\right)\frac{f(0)l(0)}{f(\widetilde{i})^2}f'(\widetilde{i}). \tag{14}$$

Proof In Appendix. ■

Condition (11) is the ex post incentive compatibility requirement for agents whose investment shock θ is realized at zero. It guarantees that low-skilled agents cannot benefit from over-reporting their skill by declaring that their θ is equal to one. Condition (12) is the first order condition from the agents' optimal deviation problem. It ensures that agents find it optimal ex ante to invest in human capital the amount recommended by the planner, i, given that they plan to truthfully announce their realization of the shock θ ex post. The role of condition (13) is to ensure that agents indeed do find it optimal to truthfully reveal their θ. This condition guarantees that utility provided to an agent who makes the recommended investment ex ante and truthfully reveals his type ex post (i.e., his realization of θ) is at least as large as the level of utility that this agent can attain by making the following joint (i.e., two-dimensional) deviation from truth-telling: lie ex post about the realization of the investment shock if $\theta = 1$ and deviate ex ante from the recommended human capital investment level i to the level of investment that maximizes the utility value of this deviation strategy. This deviation's optimal level of human capital investment is denoted by $\widetilde{i}$. The value of $\widetilde{i}$ is given implicitly by equation (14).

Note that the ex post IC constraint for the high-skilled,

$$u(c_1(1)) + v(l(1)) \geqslant u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i)}\right), \tag{15}$$

which guarantees that the high-skilled weakly prefer to report their type truthfully, does not show up in Lemma 1. This constraint is important in Mirrlees economies with exogenous skills because it binds in these environments. In our environment, this constraint does not play any role because the IC condition (13) is a tighter constraint than (15). In particular, at the optimum, (15) holds as a strict inequality. Intuitively, since (13) discourages the joint deviation of shirking and optimally deviating from the recommended investment level i, the one-dimensional deviation of just shirking is also eliminated by (13).

Lemma 1 shows that conditions (11), (12) and (13) are necessary and sufficient for incentive compatibility of an allocation. This lemma is useful because it simplifies the IC constraints. In particular, it replaces the IC conditions of Definition 3, which involve maximization problems, with two inequalities and one equality condition.

By Lemma 1, we can equivalently express the social planner's problem P1 as the problem of maximization of ex ante expected utility (1) subject to the resource feasibility constraints (RF) and the incentive compatibility constraints given by (11), (12) and (13). We thus define the following:

Problem P2

$$\begin{aligned}&\max_{(c,i,k_1,l,K,Y)} u(c_0) + \beta \sum_{\theta\in\Theta} \pi(\theta)\left[u(c_1(\theta)) + v(l(\theta))\right] \\ &s.t. \quad \text{(RF)}, (11), (12), (13).\end{aligned}$$

Problems P1 and P2 are equivalent, i.e., have the same objective and the same constraint set. The optimal allocation, therefore is given as the solution to problem P2. We characterize the optimum by studying problem P2. Lemma 2 below demonstrates that, in problem P2, the ex ante incentive constraint (13) is the only IC constraint that binds.

Lemma 2 *In problem P2, the constraint* (13) *binds, and the constraints* (11) *and* (12) *are non-binding.*

Proof In Appendix. ■

Lemma 2 shows that the deviation strategy of shirking ex post and investing $\widetilde{i}$ instead of i ex ante is in fact the best deviation strategy available to the agents in this model. This strategy generates the binding constraint (13). Lemma 2 shows that the ex post IC constraint for the low-skilled, (11), is slack at the optimum. This slackness is a standard feature of Mirrleesian models. Lemma 2 also demonstrates that the constraint (12) is not binding at the optimum. This is intuitive. Condition (12) guarantees that agents find it optimal to follow the planer's recommended human capital investment i, given that they do not plan to misrepresent ex post the realized value of the investment shock θ. But in this case, i.e., when ex post incentives are taken care of, the planner and an agent who considers a deviation from i maximize the same, undistorted by ex post misrepresentations, expected utility. Therefore, the same value of i maximizes both the planner's and the agent's objective and, in effect, (12) is always satisfied at the solution to P2, and it never binds in this problem.

To summarize, we have shown in Lemma 1 that the optimal allocation is a solution to problem P2. Lemmas 2 demonstrates that constraints (11) and (12) do not bind while (13) does bind in P2. Therefore, the optimal allocation is a solution to the problem of maximization of ex ante expected utility (1) subject to the resource feasibility constraints and a single IC equality constraint, namely, the ex ante incentive compatibility constraint (13) satisfied as equality.

The first order (FO) conditions of this problem together with the RF constraints and the binding IC condition completely determine the optimal allocation. In the next subsection, we use these conditions to demonstrate some properties of the optimal allocation.

3.2 Properties of the optimal allocation

Directly from the FO conditions of the planner's problem, we obtain some of the usual properties that optimal allocations possess in private information Mirrlees economies with exogenous skills:

$$c_1^*(0) < c_1^*(1), \tag{16}$$
$$f(0)l^*(0) < f(i^*)l^*(1), \tag{17}$$
$$u'(c_1^*(0)) > \frac{-v'(l^*(0))}{f(0)F_2(K^*,Y^*)}, \tag{18}$$
$$u'(c_1^*(1)) = \frac{-v'(l^*(1))}{f(1)F_2(K^*,Y^*)}, \tag{19}$$

where K^* and Y^* denote the optimal aggregate physical capital K, and aggregate effective labor Y, respectively.

Inequalities (16) and (17) show that the optimal allocation features less than full consumption insurance and a larger effective labor assignment to the more productive agents. Also, (18) shows the optimality of the so-called *intratemporal wedge* between the low-skilled type's marginal utility of an additional unit of consumption and this type's marginal cost of producing this additional unit. As (19) indicates, no such wedge is optimal for the high-skilled (no distortion at the top).

Another important property of optimal allocations in Mirrlees economies is the so-called reciprocal or inverse Euler condition. This condition was derived by Rogerson (1985) in a repeated moral hazard model and extended by Golosov, Kocherlakota and Tsyvinski (2003) and Kocherlakota (2005) to a large class of dynamic Mirrlees environments in which agents' IMRS are public. Adapted to the notation of our model, the Rogerson condition is given as follows:

$$u'(c_0) = \frac{\beta r}{E[1/u'(c_1)]}, \tag{20}$$

where r denotes the marginal return on physical capital investment, i.e., $r = F_1(K,Y) = Z_1(K,Y) + 1 - \delta$.

Proposition 1 below demonstrates that in our environment the optimal allocation does not satisfy the Rogerson's reciprocal Euler equation. Instead, it does satisfy an analog of this condition, which we term "modified Rogerson condition".

Proposition 1 *At the optimal allocation* $(c^*, i^*, k_1^*, l^*, K^*, Y^*)$,

$$u'(c_0^*) = \frac{\beta r^*}{E[1/u'(c_1^*)]} - \alpha\left(u'(c_0^*) - u'(c_0^* + i^* - \widetilde{i^*})\right), \tag{21}$$

where

$$\alpha = \frac{\frac{u'(c_1^*(0))}{u'(c_1^*(1))} - 1}{1 + \frac{\pi(1)}{\pi(0)}\frac{u'(c_1^*(0))}{u'(c_1^*(1))}}. \tag{22}$$

Proof In Appendix. ■

A simple but important property of the optimal allocation follows from (12), (14) and (17):

$$\widetilde{i^*} < i^*. \tag{23}$$

Inequality (23) shows that, at the optimum, agents' best deviation strategy involves an under-investment in human capital and over-consumption at $t = 0$. This fact has significant implications for the structure of the optimal allocation. Using (23) and (16), we get that the term subtracted on the right-hand side of (21) is strictly positive. Therefore, our modified Rogerson equation (21) implies that

$$u'(c_0^*) < \frac{\beta r^*}{E[1/u'(c_1^*)]}, \tag{24}$$

which demonstrates that the optimal allocation of our environment does not satisfy the Rogerson condition (20). It follows from (24), combined with the Jensen inequality, that, similar to Mirrlees economies with exogenous skills, the so-called *intertemporal wedge* is optimal in our environment:

$$u'(c_0^*) < r^* \beta E\left[u'(c_1^*)\right].$$

In a sense to be made precise in Section 5, this wedge is larger in our endogenous skill environment than in standard Mirrlees environments with exogenous skills. We show in Section 5 that this larger intertemporal wedge translates into a larger volatility of marginal capital tax rates.

In order to explain the difference between the usual Rogerson condition (20) of the standard Mirrlees model and the modified Rogerson condition (21) of our model, we first provide a variational argument for and present the intuition behind the usual Rogerson equation (20). Then we extend this argument and intuition to our model. To this end, suppose, in the context of our model, that human capital investment i is publicly observable, which in terms of incentives is equivalent to the standard Mirrlees economy in which skills are given exogenously. Since the initial capital endowment k_0, and physical capital investment k_1 are observable, consumption c_0 can be now perfectly inferred from k_0 and the observations of k_1 and i. Thus, the IC constraints of Lemma 1 reduce to the following two ex post IC constraints:

$$u(c_1(0)) + v(l(0)) \geq u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(0)}\right), \tag{25}$$

$$u(c_1(1)) + v(l(1)) \geq u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i)}\right). \tag{26}$$

We know from the standard Mirrlees model that the first of those constraints is slack at the optimum. Let us then disregard it for the purpose of this discussion. The incentive problem in this environment is to discourage the high-skilled from pretending to be low-skilled (all agents make the recommended human capital investment i because this investment is now observable and any deviation can be severely punished). The deviation strategy that makes (26) bind in the planner's problem of this environment consists of declaring the low realization of the investment shock θ when it really is high, which results in the low effective labor assignment $f(0)l(0)$ in place of the high assignment $f(i)l(1)$ (shirking). This strategy provides the deviator with the total ex ante utility of

$$v^{dev} = u(c_0) + \beta\pi(0)\left(u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(0)}\right)\right) + \beta\pi(1)\left(u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i)}\right)\right),$$

while the truth-teller's utility is equal to

$$v^{tt} = u(c_0) + \beta\pi(0)\left(u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(0)}\right)\right) + \beta\pi(1)\left(u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i)}\right)\right).$$

The IC condition (26) is satisfied if and only if $v^{tt} \geq v^{dev}$.

Take now an incentive compatible and resource feasible allocation (c, i, k_1, l, K, Y) and consider a small re-allocation of consumption between date 0 and date 1 via a change in physical capital investment dk_1. For concreteness, take $dk_1 > 0$. This investment hike reduces the date 0 consumption c_0 by dk_1 and makes extra resources available at date 1 in the amount of rdk_1. The magnitude of the change in the social welfare that results from this small increase in physical capital investment depends in general on how those extra resources available at date 1 are distributed among the two ex post types of agents. The key point here is that this distribution has to preserve incentives. Since the initial allocation (c, i, k_1, l, K, Y) is incentive compatible, incentives will be preserved if the gain in the utility of the deviator, v^{dev}, is not larger than the gain in the utility of the truth-teller, v^{tt}. The cut in the ex ante consumption c_0 affects the values of the deviation strategy and the truth-telling strategy identically: the values of both strategies are decreased by $u'(c_0)dk_1$. The ex post consumption utility gains of the two strategies, however, depend on how the extra resources are split between those who declare $\theta = 0$ and those who declare $\theta = 1$. In the relevant case of less than full insurance of the consumption risk, we have $u'(c_1(0)) > u'(c_1(1))$ and thus the gain in the social welfare would be maximized if a larger part of rdk_1 were allocated to those who declare $\theta = 0$. This, however, would in general violate incentives by giving a larger gain in the utility of consumption at date 1 to the deviator, as the deviator always declares $\theta = 0$. Therefore, the best that the planner can do ex post is to distribute the resources rdk_1 in such a way that the gain in the ex post utility of consumption is the same for all values of θ. Denoting the amount allocated to those who declare θ by $\rho(\theta)$, we therefore need to have

$$u'(c_1(\theta))\rho(\theta) = \triangle, \tag{27}$$

with the same $\triangle$ for all θ, and with

$$\sum_{\theta \in \Theta} \pi(\theta)\rho(\theta) = rdk_1. \tag{28}$$

Note that ρ is low when marginal utility of consumption is high, which is inefficient in the "first-best" sense. The total gain in the ex ante expected utility, under truth-telling, that results from this distribution of the surplus is equal to

$$\sum_{\theta \in \Theta} \pi(\theta)\beta u'(c_1(\theta))\rho(\theta) = \beta\triangle.$$

Substituting (27) into (28) and solving for $\triangle$, we get

$$\triangle = \frac{rdk_1}{\sum_{\theta \in \Theta} \frac{\pi(\theta)}{u'(c_1(\theta))}} = \frac{r}{E[1/u'(c_1)]} dk_1.$$

Adding up the loss due to decreased consumption at date 0 and the gain due to increased consumption at date 1, we get that the total change in the ex ante expected utility, under truth-telling (i.e., on the equilibrium path), is given by

$$\left(-u'(c_0) + \frac{\beta r}{E[1/u'(c_1)]}\right) dk_1.$$

If the initial allocation (c, i, k_1, l, K, Y) is optimal, the above expression has to be equal zero for all dk_1, which happens only if the expression in the bracket is zero, which yields the Rogerson condition (20).

Let us now return to our environment, in which human capital investment i and consumption c_0 are indistinguishable to the planner, and thus private information of the agent. A reexamination of the foregoing variational argument shows why the Rogerson condition (20) does not characterize the optimum in our environment. We know from Lemma 2 that the only IC constraint that binds in the planner's problem P2 is the constraint (13). The incentive problem in our environment, then, is to discourage the agents from doing a joint deviation of under-investing in human capital ex ante and shirking ex post. At this deviation strategy, the deviator invests $\widetilde{i} < i$ and shirks. A key difference between the best deviation strategy in the observable (or exogenous) human capital investment model and our model with private human capital investment is that date-zero consumption of the deviator is not equal to the date zero consumption of the truth-teller in our model. The value of the deviation strategy in our private human capital investment model is

$$v^{dev} = u(c_0 + i - \widetilde{i}) + \beta\pi(0)\left(u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(0)}\right)\right) + \beta\pi(1)\left(u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(\widetilde{i})}\right)\right),$$

where $\widetilde{i}$ is given as a solution to (14). The utility of an agent who follows the planner's recommendation at both dates is equal to

$$v^{tt} = u(c_0) + \beta\pi(0)\left(u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(0)}\right)\right) + \beta\pi(1)\left(u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i)}\right)\right).$$

The IC condition (13) is satisfied if and only if $v^{tt} \geq v^{dev}$.

As before, take now an incentive compatible and resource feasible allocation (c, i, k_1, l, K, Y) and consider a small hike in the physical capital investment dk_1 funded by a cut in date-0 recommended consumption c_0. Unlike in the previous case of observable human capital investment, the values of the deviation strategy v^{dev} and the equilibrium strategy v^{tt} are affected differently by this cut in c_0. The deviator is less hurt by the cut in c_0 than the on-equilibrium agent because the deviator's date-0 consumption $c_0 + i - \widetilde{i}$ is strictly larger than the on-equilibrium date-0 consumption c_0, and u is concave. This is why the Rogerson equation does not hold at the optimum of the private human capital investment model. Ex post, the planner cannot use the same distribution of the resource surplus that was used in the observable investment model, i.e., the distribution ρ that gives the same gain in date-1 consumption utility to both the deviator and the on-equilibrium agent. This distribution would make the deviation strategy more attractive relative to the on-equilibrium strategy. If the utility of the deviator at date 1 were to increase by as much as the utility of the on-equilibrium agent, given that the deviator loses less utility at date 0, the overall value of the deviation strategy would increase, relative to the equilibrium strategy, which would in general violate incentives.

In order to preserve incentives, therefore, the planner has to use at date-1 a distribution of the surplus that strictly favors the on-equilibrium agent over the deviator. This advantage has to be just large enough to offset the deviator's advantage that results from the equal cut in the date-0 consumption of the deviator and the on-equilibrium agent. In order to do so, this distribution of the surplus has to give even more resources to those agents who announce $\theta = 1$ ex post, as the deviator always announces $\theta = 0$. Thus, on the equilibrium path, this distribution is even more inefficient, relative to the first best, than the distribution ρ of the model with observable investment. Denote by $\eta(\theta)$ the amount by which the planner increases consumption of the agent who declares θ at data 1. In order to maintain incentives, η must be such that the overall change in v^{dev} is the same as the overall change in v^{tt}, i.e.,

$$-u'(c_0 + i - \widetilde{i})dk_1 + \beta\pi(0)u'(c_1(0))\eta(0) + \beta\pi(1)u'(c_1(0))\eta(0) = \triangle, \quad (29)$$

$$-u'(c_0)dk_1 + \beta\pi(0)u'(c_1(0))\eta(0) + \beta\pi(1)u'(c_1(1))\eta(1) = \triangle. \quad (30)$$

Using (29) and $\sum_{\theta\in\Theta}\pi(\theta)\eta(\theta) = rdk_1$ to eliminate η from (30), with a little bit of algebra, we get that the total change in the utility of the on-equilibrium agent is given by

$$\left(-u'(c_0) + \frac{\beta r}{\frac{\pi(0)}{u'(c_1(0))} + \frac{\pi(1)}{u'(c_1(1))}} - \frac{\frac{u'(c_1(0))}{u'(c_1(1))} - 1}{1 + \frac{\pi(1)}{\pi(0)}\frac{u'(c_1(0))}{u'(c_1(1))}}\left(u'(c_0) - u'(c_0 + i - \widetilde{i})\right)\right) dk_1.$$

If the initial allocation (c, i, k_1, l, K, Y) is optimal, the above expression has to equal zero for all dk_1, which means that the expression that multiplies dk_1 has to be zero, which yields the modified Rogerson equation (21).

Intuitively, as the deviation opportunities for the agents are richer in the environment with private human capital investment, the incentive cost of physical capital investment, measured by the dispersion of the ex post distribution of the return, is larger. This leads to the modified Rogerson equation (21) and inequality (24).

In the existing literature on dynamic optimal taxation, physical capital is the only asset that the planner can use to transfer resources between two points in time. In the environment we study in this paper, the planner can alternatively use human capital as an indirect vehicle for moving consumption between dates 0 and 1.

Let us denote by R the human capital investment's marginal return $\pi(1)f'(i)l(1)F_2(K, Y)$. Proposition 2 below derives a version of the modified Rogerson equation that applies to this return at the optimum.

Proposition 2 *At the optimal allocation* $(c^*, i^*, k_1^*, l^*, K^*, Y^*)$,

$$u'(c_0^*) = \frac{\beta R^*}{E[1/u'(c_1^*)]} - \alpha u'(c_0^*), \quad (31)$$

where, as before, α *is given in (22).*

Proof In Appendix. ■

The reason for the difference between the modified Rogerson equation for physical capital investment (21) and human capital investment (31) can be seen from the variational argument we have discussed before. If the planner increases physical capital investment by dk_1 and decreases the recommended consumption c_0 by the same amount, then, as we discuss above, the date-0 utility of an agent who follows the best deviation strategy decreases by $-u'(c_0+i-\widetilde{i})dc_0 = u'(c_0+i-\widetilde{i})dk_1$, which is less than $u'(c_0)dk_1$ but more than 0. If, in turn, the planner increases human capital investment by di and decreases the recommended consumption c_0 by the same amount, the date-0 consumption of the agent who follows the deviation strategy does not drop at all. To see this note that the date-0 utility of the deviating agent decreases by $-u'(c_0+i-\widetilde{i})dc_0 + u'(c_0+i-\widetilde{i})di = 0$. Thus, a marginal increase in the recommended human capital investment increases the value of the deviation strategy even more rapidly than a marginal increase in the physical capital investment does. In order to compensate for it, the planner has to split the date-1 return on this investment, Rdi, with an even larger bias, relative to the distribution η, toward those agents who announce $\theta = 1$, which decreases the social value of human capital investment.

This additional relative inefficiency of the best incentive feasible distribution of the return on human capital investment makes the social return on a unit of human capital investment lower than the social return on a unit of physical capital investment. At the optimum, the social rates of return on those two assets have to be equal. This means that the physical rate of return on the human capital investment, R, must exceed the physical rate of return on physical capital investment, r, at the optimum. This fact is highlighted in the following corollary.

Corollary 1 *At the optimal allocation* $(c^*, i^*, k_1^*, l^*, K^*, Y^*)$,

$$R^* > r^*. \tag{32}$$

Proof Follows directly from (21) and (31). ■

The inequality (32) establishes the optimality of a wedge between the marginal rates of return to physical and human capital investment. We will refer to this wedge as the *asset return wedge*. In Section 6, we study numerically the properties of this wedge.

4 Decentralization

Following the Mirrlees approach to optimal taxation, we study the question of decentralization of the optimal allocation in an equilibrium with taxes, in which the only constraints on the set of tax instruments available to the government are imposed by the informational frictions primitive to the economic environment studied.[6] In general, many such decentralizations are possible. Following Kocherlakota (2005), we confine our attention to decentralizations in which the tax on capital is linear in capital and labor taxes depend solely on labor income. In Proposition 3, we show the existence of such a decentralization and provide a complete characterization of the optimal tax system.

In Proposition 4, we show that, as in Kocherlakota (2005), the expected value of the marginal capital tax rate is zero. It follows that government's revenue from capital taxation is optimally equal to zero. The proof of this result given in Kocherlakota (2005) relies on the fact that the optimal allocation of his environment satisfies a generalized version of the first-order condition of Rogerson (1985). By Proposition 1, the Rogerson condition does not hold in our environment. Thus, the existence of a decentralization with linear capital taxes and the zero average tax result that we

[6] See Kocherlakota (2005a) for a recent review of this literature.

obtain are a generalization of the results of Kocherlakota beyond the class of economies in which the IMRS are public and the Rogerson condition holds.

4.1 Equilibrium

We now introduce a set of markets in which agents trade taking prices and taxes as given. There is spot trade at the ex post date. Consumption, human capital investment, and savings decisions are made ex ante. Agents save by accumulating physical capital, i.e., there are no finanical assets. Our results on optimal taxes do not depend on this restriction. As in Kocherlakota (2005), the introduction of a bond into our model would not change anything as long as all wealth receives the same tax treatment.

There is a single representative firm that operates the technology of production ex post. Due to free entry and constant returns to scale, the single-firm assumption is without a loss of generality. The firm employs physical capital and effective labor at prices r and w, respectively. The firm takes these prices as given. In equilibrium, inputs are paid their marginal product.

At the ex ante date, agents decide how to allocate their initial endowment of capital k_0 between consumption c_0, physical capital investment k_1, and human capital investment i. At the ex post date, agents supply their capital k_1 and effective labor $y = f(\theta i)l(\theta)$ to the firm.

Publicly observable are capital holdings k_1 and effective labor y. Since there is no investment in human capital at time $t = 1$, the value of private consumption c_1 can be inferred from the observables. Human capital investment i and consumption c_0, the stock of human capital $h = f(\theta i)$, and the amount worked $l(\theta)$ are private.

We follow Kocherlakota (2005) in restricting attention to tax systems consisting of a linear tax $\widetilde{\tau}$ on capital and a possibly non-linear tax on effective labor $\widetilde{\phi}$. Formally, the tax system consists of two functions $\widetilde{\tau}(y)$ and $\widetilde{\phi}(y)$, where $\widetilde{\tau}$ is the marginal capital tax rate, and $\widetilde{\phi}$ is the amount of the labor income tax. Both the labor tax and the marginal capital tax rate are functions of the observable effective labor provided by the agent, which makes this tax system linear in the physical capital holdings k_1.

Definition 6 *Given a tax system $(\widetilde{\tau}, \widetilde{\phi})$, a competitive equilibrium is an allocation of the consumption good, physical capital, investment in human capital and labor input for the agent (c^e, k_1^e, i^e, l^e), aggregate capital stock and effective labor input (K^e, Y^e), and prices (r, w) such that*

1. *given taxes $(\widetilde{\tau}, \widetilde{\phi})$ and prices (r, w), (c^e, k_1^e, i^e, l^e) solves*

$$\max_{c_0, c_1(\theta), k_1, i, l(\theta)} u(c_0) + \beta \sum_{\theta} \pi(\theta)[u(c_1(\theta)) + v(l(\theta))]$$

   subject to

$$c_0 + i + k_1 \leq k_0,$$
$$c_1(\theta) \leq (1 - \widetilde{\tau}(f(\theta i)l(\theta)))rk_1 + wf(\theta i)l(\theta) - \widetilde{\phi}(f(\theta i)l(\theta)),$$

$$c_0, c_1(\theta), i, k_1, l(\theta) \geq 0, \text{ for } \theta \in \Theta, \; k_0 > 0 \text{ given,}$$

2. *prices (r, w) are given by*

$$
\begin{aligned}
r &= F_1\left(K^e, Y^e\right), \\
w &= F_2\left(K^e, Y^e\right),
\end{aligned}
$$

3. markets clear

$$
\sum_{\theta} \pi(\theta) c_1^e(\theta) + G = F\left(K^e, Y^e\right),
$$

$$
K^e = k_1^e,
$$

$$
Y^e = \sum_{\theta \in \Theta} \pi(\theta) f\left(\theta i^e\right) l^e(\theta).
$$

In the above definition, the budget constraints in the agents' problem incorporate effective-labor-dependent taxes. The marginal factor pricing equations are equivalent to firm profit maximization under perfect competition in the production sector with constant returns to scale. Since profits must be zero at any equilibrium, we do not need to specify the ownership of the firm. The agent's budget constraints and the market clearing conditions imply that the government budget is balanced.

4.2 Equilibrium with a simplified tax system

As we allow the tax schedule to be an arbitrary function of effective labor, it can be chosen such that it effectively limits the choice of the effective labor by the agent to $\{f(0)l^*(0), f(i^*)l^*(1)\}$. At the optimal allocation, effective labor of agents whose productivity shock is $\theta = 0$ is $y = f(0)l^*(0)$, and effective labor provided by agents whose productivity shock is $\theta = 1$ is $y = f(i^*)l^*(1)$. The tax system is to be designed so as to implement this behavior in equilibrium. It is then easy to discourage all kinds of behavior inconsistent with the optimal allocation and such that the effective labor levels are different from $f(0)l^*(0)$ or $f(i^*)l^*(1)$. Namely, if the observed effective labor $y \notin \{f(0)l^*(0), f(i^*)l^*(1)\}$, then the punishment the tax code inflict upon the agent can be arbitrarily severe. For example, we can set the tax rate at 100% on both labor income and capital. Therefore, we set

$$
\tilde{\phi}(y) = wy, \tag{33}
$$

$$
\tilde{\tau}(y) = 1, \tag{34}
$$

for $y \notin \{f(0)l^*(0), f(i^*)l^*(1)\}$.

The problem of the choice of taxes for $y \in \{f(0)l^*(0), f(i^*)l^*(1)\}$ is less straightforward. If the observed effective labor $y = f(\theta i)l(\theta)$ is equal to $f(0)l^*(0)$, it is not publicly known if this is due to a low productivity shock $\theta = 0$ or different than recommended choice of investment i and labor l. If the observed effective labor $y = f(\theta i)l(\theta)$ is equal to $f(i^*)l^*(1)$, then, again, it is not publicly known if this is due to a high productivity shock $\theta = 1$ or different than recommended choice of investment and labor. Taxes $\widetilde{\phi}(y)$ and $\widetilde{\tau}(y)$ for $y \in \{f(0)l^*(0), f(i^*)l^*(1)\}$ must be set so as to induce agents to invest the optimal amount i^* and provide labor $l^*(0)$ ex post if $\theta = 0$ and $l^*(1)$ ex post if $\theta = 1$. In what follows, we introduce the following notation:

$$\phi(\theta) \quad : \quad = \widetilde{\phi}(f(\theta i^*) l^*(\theta)), \tag{35}$$

$$\tau(\theta) \quad : \quad = \widetilde{\tau}(f(\theta i^*) l^*(\theta)), \tag{36}$$

where $\phi(\theta)$ and $\tau(\theta)$ are, respectively, the labor income and capital tax for agents whose observed effective labor is consistent with optimal behavior, given the shock realization of θ.

The simplified tax schedule, therefore, consists of four numbers $\phi(0)$, $\phi(1)$, $\tau(0)$, $\tau(1)$, with taxes for $y \notin \{f(0)l^*(0), f(i^*)l^*(1)\}$ given in (33) and (34).

Let us now examine the agent's utility maximization problem under the simplified tax system. At $t = 0$, agents choose c_0, i, and k_1 subject to the budget constraint $c_0 + i + k_1 \leq k_0$. At $t = 1$, agents rent their capital to the firm and choose their labor $l(\theta)$. If, given i chosen at $t = 0$ and θ realized at the beginning of period $t = 2$, the choice of labor $l(\theta)$ is such that $f(\theta i)l(\theta) \notin \{f(0)l^*(0), f(i^*)l^*(1)\}$, then labor income and capital taxes are 100%, which leads to zero consumption. No such choice can be optimal for the agent. Effectively, then, given i and θ, the choice of labor $l(\theta)$ is restricted to such values that effective labor $f(\theta i)l(\theta)$ belongs to $\{f(0)l^*(0), f(i^*)l^*(1)\}$, i.e., $l(\theta) = \frac{f(0)l^*(0)}{f(\theta i)}$ or $l(\theta) = \frac{f(i^*)l^*(1)}{f(\theta i)}$. In the first case, taxes due are given by $\tau(0)rk_1$ and $\phi(0)$. In the second case, taxes due are given by $\tau(1)rk_1$ and $\phi(1)$.

To write the agent's problem under this simplified tax system formally, we define a choice variable $s(\theta) \in \{0, 1\}$ that indicates the choice to provide the high amount of effective labor, $f(i^*)l^*(1)$, conditional on θ. I.e., $s(\theta) = 1$ represents the decision to provide $f(i^*)l^*(1)$ units of effective labor in state θ, and the choice $s(\theta) = 0$ represents the decision to provide $f(0)l^*(0)$ units of effective labor in state θ.

Using this notation, we can express agents' utility maximization problem under a simplified tax system ϕ, τ as follows. Given taxes (τ, ϕ) and prices (r, w), agents solve

$$\max_{c_0, c_1(\theta), k_1, i, s(\theta)} u(c_0) + \beta \sum_{\theta} \pi(\theta) \left\{ u(c_1(\theta)) + [1 - s(\theta)]\, v\left(\frac{f(0)l^*(0)}{f(\theta i)}\right) + s(\theta) v\left(\frac{f(i^*)l^*(1)}{f(\theta i)}\right) \right\}$$

subject to

$$c_0 + i + k_1 \leq k_0,$$

$$c_1(\theta) \leq (1 - s(\theta))\left[(1 - \tau(0))rk_1 + wf(0)l^*(0) - \phi(0)\right] + s(\theta)\left[(1 - \tau(1))rk_1 + wf(i^*)l^*(1) - \phi(1)\right],$$

$$c_0, c_1(\theta), i, k_1 \geq 0 \text{ for } \theta \in \Theta.$$

Note in the above problem that the choices of i and $s(\theta)$ determine the level of labor $l(\theta)$, so $l(\theta)$ is not formally included as a choice variable.

The definition of competitive equilibrium under simplified taxes (τ, ϕ) is identical to Definition 6, but agents' utility maximization problem is replaced by the above-defined agents' problem with simplified taxes.

4.3 Optimal taxes

Definition 7 *A tax system is optimal if, under this system, there exists a competitive equilibrium such that the equilibrium allocation* $(c^e, k_1^e, i^e, l^e, K^e, Y^e)$ *coincides with the optimal allocation* $(c^*, k_1^*, i^*, l^*, K^*, Y^*)$.

Proposition 3 establishes the existence of an optimal simplified tax system.

Proposition 3 *In the class of simplified tax systems with linear capital taxes, there exists an optimal tax system. In particular, the following simplified tax system* (τ^*, ϕ^*) *is optimal:*

$$\tau^*(0) = 1 - \frac{u'(c_0^* + i^* - \widetilde{i^*})}{r^* \beta u'(c_1^*(0))}, \tag{37}$$

$$\tau^*(1) = 1 - \frac{u'(c_0^*) - \pi(0) u'(c_0^* + i^* - \widetilde{i^*})}{\pi(1) r^* \beta u'(c_1^*(1))}, \tag{38}$$

$$\phi^*(\theta) = (1 - \tau^*(\theta)) r^* k_1^* + f(\theta i^*) l^*(1) w^* - c_1^*(\theta) \ \text{ for } \theta \in \Theta, \tag{39}$$

where

$$r^* = F_1\left(k_1^*, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i^*) \, l^*(\theta)\right),$$

$$w^* = F_2\left(k_1^*, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i^*) \, l^*(\theta)\right),$$

and where $\widetilde{i^*}$ *is the value of* i *that solves*

$$u'(c_0^* + i^* - i) = -\pi(1) \beta v'\left(\frac{f(0) l^*(0)}{f(i)}\right) \frac{f(0) l^*(0)}{f(i)^2} f'(i).$$

Proof In Appendix. ■

The proof of this proposition is constructive. We show that with capital and labor prices $(r, w) = (r^*, w^*)$ and with taxes (τ^*, ϕ^*) given in (37)–(39), agents choose the optimal allocation in equilibrium. The proof amounts to checking that the optimal allocation satisfies the equilibrium conditions of Definition 6. Market clearing and competitive pricing conditions are immediate. The main part of the proof shows that agents' equilibrium choices of effective labor, consumption, human and physical capital investment are consistent with the optimal allocation.

Essentially, our model in which agents trade in a set of markets subject to prices and taxes constitutes a mechanism in the sense of mechanism design theory. The optimal allocation is an equilibrium outcome of a different mechanism, the direct revelation mechanism. Our decentralization exercise amounts to checking that, in the class of environments studied, there exists an equilibrium outcome of our market/tax mechanism that coincides with the desired equilibrium outcome of the direct revelation mechanism, i.e., the optimal allocation.

Following Kocherlakota (2005), in our market/tax mechanism we use taxes that are arbitrarily nonlinear in effective labor and set them at 100% of all wealth if an agent's observed effective labor does not equal to effective labor assigned at the optimum to some type of agent. As in Kocherlakota, this guarantees that, in the market/tax equilibrium, agents whose type is θ choose to supply effective labor in an amount that in the optimal allocation is assigned to an agent whose type is $\hat{\theta}$. The last, and most important, step of the proof involves showing that indeed $\hat{\theta} = \theta$, and that, conditional on this choice of effective labor in a market/tax equilibrium, the choices of consumption, human and physical capital investment are consistent with the optimal allocation.

In Kocherlakota's (2005) environment, this last step of the proof is broken into two sub-steps. First, Kocherlakota uses the Rogerson equation, which holds in his environment with public IMRS, to show that conditional on the effective labor choice of type $\hat{\theta}$, the agents' choices of the remaining decision variables in the market/tax mechanism coincide with the values that those variables are assigned in the direct revelation mechanism to an agent who declares to be of type $\hat{\theta}$. I.e., if an agent chooses in the market/tax equilibrium of Kocherlakota to provide effective labor that in the direct revelation equilibrium is assigned to an agent who declares to be of type $\hat{\theta}$, then this agent will also choose in the market/tax equilibrium the same path of capital accumulation and consumption that the planner assigns at the optimal allocation to an agent who declares to be of type $\hat{\theta}$. Therefore, for any effective labor strategy $\hat{\theta}$, the utility that an agent attains in the market/tax equilibrium is equal to the utility that the optimal allocation provides to an agent who declares type $\hat{\theta}$. Using this type-by-type equality of utility provided to agents in the two mechanisms, the second sub-step of Kocherlakota's proof simply invokes incentive compatibility of the optimal allocation to demonstrate that, in the equilibrium of the market/tax mechanism, agents of type θ choose the effective labor, and therefore the rest of the allocation too, that is socially optimal for an agent of type θ, which completes the proof.

Our proof of Proposition 3 follows Kocherlakota's proof of decentralization with one important difference. The last step of our proof cannot be split into the two sub-steps that Kocherlakota uses. In particular, it is not true in our model that in the market/tax equilibrium consumption and investment choices of an agent who provides effective labor optimal for type $\hat{\theta}$ coincide with the consumption and investment levels that are assigned to type $\hat{\theta}$ at the optimal allocation. Under taxes (τ^*, ϕ^*) given in (37)–(39) and prices (r^*, w^*), some off-equilibrium joint effective labor, consumption and investment deviation strategies attain utility levels strictly higher than the utility levels that those same off-equilibrium effective labor strategies attain in the direct revelation equilibrium, i.e., at the optimal allocation. Therefore, the incentive compatibility of the optimal allocation in itself does not guarantee, in general, the incentive compatability of the optimal allocation in the market/tax mechanism. Despite that, in our environment, we are able to prove that the values attained by those joint deviations do not exceed the value of the optimum.[7]

The optimal taxes (τ^*, ϕ^*) are chosen in our model so as to ensure that the logic of Kocherlakota's proof applies to the truth-telling effective labor strategy and the off-equilibrium effective labor strategy for which the IC constraint binds in the planner's problem. I.e., taxes are chosen in such a way that, conditionally on those effective labor strategies, agents' market/tax equilibrium consumption and investment choices do coincide with the quantities optimally assigned by the planner. The other effective labor deviation strategies do not bind in the planner's problem, so the utility levels they provide are strictly less than the utility of the truth-telling strategy in the direct

[7] In the present version of our proof, for some parameter values, in order to reach this conclusion we need to make an additional assumption about the behavior of endogenous variables. Numerous numerical experiments we conducted strongly suggest that this assumption follows from the primitives of our model, but, at this point, we don't have an analytical proof to support this conjecture. As our proof of Proposition 3 makes apparent, however, without a need for that additional assumption we do obtain our decentralization result for an open subset of economies, which makes our result generic. In particular, we have an analytical proof that decentralization obtains in all economies in which the value of $f(0)$ is sufficiently small.

revelation mechanism. In our market/tax mechanism, by simultaneously adjusting consumption and investment decisions, agents are able to increase the utility values attained by these effective labor strategies above the levels that these strategies provide in the direct revelation mechanism, and thus Kocherlakota's proof does not apply. It turns out however that, for any such off-equilibrium effective labor strategy, the gain in utility produced by the joint deviation of investment and consumption is strictly less than the amount by which the utility of truth-telling exceeds the utility of this particular effective labor strategy in the direct revelation mechanism. Thus, although these off-equilibrium strategies do strictly better under the market/tax mechanism than under the direct revelation mechanism, the improvement is never sufficient to exceed the value the effective labor strategy associated with truth-telling and, thus, the market/tax mechanism, under taxes (τ^*, ϕ^*) in (37)–(39), remains incentive compatible.

In our model, we use the same market/tax mechanism with linear capital taxes that is used in Kocherlakota (2005). In the model of Kocherlakota (2005), however, any off-equilibrium strategy generates the same utility both in the market/tax mechanism and in the direct revelation mechanism. In our model, some off-equilibrium strategies are strictly more attractive in the market/tax mechanism relative to the direct revelation mechanism. Therefore, the linear capital tax mechanism in our environment is a significantly stronger relaxation of the direct revelation mechanism than it is in the environment studied in Kocherlakota (2005). The fact that this mechanism continues to work as a decentralization of the optimum in our model demonstrates robustness of the linear capital tax mechanism.

Proposition 4 below shows that Kocherlakota's zero expected wealth tax result continues to hold, despite the fact that the Rogerson condition does not hold in our environment.

Proposition 4 *At the optimal tax system* (τ^*, ϕ^*),

$$\sum_{\theta} \pi(\theta)\tau^*(\theta) = 0,$$

and

$$\tau^*(0) > 0 > \tau^*(1).$$

Proof In Appendix. ■

Proposition 4 shows that the zero expected optimal wealth tax result of Kocherlakota (2005) extends beyond the class of environments in which intertemporal marginal rates of substitution are private and individual optimal consumption allocation satisfies the Rogerson condition. Kocherlakota's proof of the zero expected tax result works off the Rogerson condition. Also, Farhi and Werning (2005) obtain the optimality of strictly positive expected wealth taxes in an environment in which the Rogerson condition does not hold. We use our modified Rogerson condition (21) to show that excepted marginal tax rate is zero in our model.

Unlike in Kocherlakota's decentralization, however, it is not true in our model that agents' conditional after-tax marginal rates of substitution are independent of idiosyncratic uncertainty and equal to the social discount factor. To see this, note that

$$
\begin{aligned}
(1-\tau^*(0))\frac{\beta u'(c_1^*(0))}{u'(c_0^*)} &= \frac{u'(c_0^*+i^*-\widetilde{i}^*)}{r^*\beta u'(c_1^*(0))}\frac{\beta u'(c_1^*(0))}{u'(c_0^*)} \\
&= \frac{1}{r^*}\frac{u'(c_0^*+i^*-\widetilde{i}^*)}{u'(c_0^*)} \\
&< \frac{1}{r^*}, \qquad (40)
\end{aligned}
$$

where the inequality follows from the fact that $i^* > \widetilde{i}^*$. Of course, we have

$$E\left[(1-\tau^*)\frac{\beta u'(c_1^*)}{u'(c_0^*)}\right] = \frac{1}{r^*}$$

because we decentralize the optimal allocation, and thus equilibrium pricing of physical capital must be consistent with the socially optimal valuation. In Kocherlakota's decentralization, however, the after-tax marginal rate of substitution $(1-\tau^*(\theta))\beta u'(c_1^*(\theta))/u'(c_0^*)$ is equal to the social discount factor $1/r^*$ for all θ. In our decentralization, as (40) demonstrates, the after-tax marginal rate of substitution varies with θ. In particular,

$$(1-\tau^*(1))\frac{\beta u'(c_1^*(1))}{u'(c_0^*)} > \frac{1}{r^*} > (1-\tau^*(0))\frac{\beta u'(c_1^*(0))}{u'(c_0^*)}.$$

Also, similar to Kocherlakota (2005), the marginal tax rate on capital is strictly positive for agents whose productivity is low and strictly negative for those whose productivity is high. The high wealth tax rate on the low skilled is needed so as to deter agents from making a joint deviation of under-investing in human capital, over-saving, and providing low effective labor independently of the realization of θ. This need to deter shirking is also present in the exogenous skill environment of Kocherlakota (2005). There, however, the joint deviation has only two dimensions. It consists of over-saving and providing low effective labor for any realization of the productivity shock. In our environment in which skills are endogenous, the double deviation of Kocherlakota turns into a triple deviation, with under-investment in human capital being the third dimension.

5 Volatility of optimal wealth taxes: endogenous versus exogenous skills

In this section, we show that skill endogenity leads to more volatility of marginal capital tax rates. More precisely, we show that it takes more volatility of marginal capital tax rates to decentralize a given consumption allocation in our environment with endogenous skills than in an exogenous-skill version of our model.

If skills were given exogenously, our model would be a special case of the model of Kocherlakota (2005). Then, the optimal marginal capital tax rates, as derived in Kocherlakota (2005), would be as follows

$$\widehat{\tau}(\theta) = 1 - \frac{u'(\widehat{c}_0)}{\widehat{r}\beta u'(\widehat{c}_1(\theta))}, \qquad (41)$$

for $\theta \in \Theta$, where hats over τ, c_0, c_1 and r indicate that these values of taxes, consumption, and the rental rate are optimal in the model with exogenous skills. Also, it follows from Kocherlakota (2005) that, with exogenous skills, the optimal consumption and rental rate satisfy the usual Rogerson condition

$$u'(\hat{c}_0) = \frac{\hat{r}}{E\left[1/\beta u'(\hat{c}_1)\right]}. \tag{42}$$

Proposition 5 *Take an endogenous-skill economy and an exogenous-skill Mirrlees economy with the same preferences over consumption. Suppose that the same consumption allocation is optimal in both economies, i.e., $(\hat{c}_0, \hat{c}_1) = (c_0^*, c_1^*)$. Then,*

$$Var(\tau^*) > Var(\hat{\tau}).$$

Proof In Appendix. ■

It is known from the work of Albanesi and Sleet (2005) and Kocherlakota (2005) that, in a market/tax decentralization of the optimum, the intertemporal wedge is implemented through the negative covariance of the marginal utility of consumption and the after-tax return on capital investment. From the point of view of an individual agent, this negative covariance makes savings a risky investment. This risk discourages agents from over-saving, which implements the intertemporal wedge.

The proof of Proposition 5 follows from the fact that the incentive problem is more severe in the environment with endogenous skills, which creates a larger social cost of physical capital investment. At a given optimal allocation, this larger cost must be offset by a larger rate of return, which implies that $r^* > \hat{r}$. This difference in the optimal returns implies a larger intertemporal wedge in the model with endogenous skills. To see this, note that at a given optimal allocation $(\hat{c}_0, \hat{c}_1) = (c_0^*, c_1^*)$, the difference in the two wedges is proportional to the difference in the optimal rates of return on capital:

$$\begin{aligned} & r^*\beta E\left[u'(c_1^*)\right] - u'(c_0^*) - \left(\hat{r}\beta E\left[u'(\hat{c}_1)\right] - u'(\hat{c}_0)\right) \\ = \;& (r^* - \hat{r})\beta E\left[u'(c_1^*)\right]. \end{aligned}$$

As the expected optimal marginal capital taxes are zero in both the endogenous and exogenous skill models, the implementation of a larger wedge requires more risk in the after-tax return on savings, i.e., more volatility of the marginal capital tax rate τ.

6 Numerical examples

In this section, we study numerically how the asset wedge depends on the degree of risk aversion of the agents, and how private information affects the optimal allocation, wedges, and taxes.

6.1 Asset Wedge

In this subsection, we study numerically the relationship between the coefficient of relative risk aversion and the asset retur wedge expressed as the ratio of the optimal marginal return to human capital investment to the optimal marginal return to physical capital investment. We set $\beta = 0.9$; $u(c) = \frac{1}{1-\sigma}c^{(1-\sigma)}$ (CRRA preferences); $v(l) = -(l)^2$; $f(h) = 0.2 + h^\gamma$ with $\gamma = 0.5$; $F(K,Y) = AK^\alpha Y^{1-\alpha} + (1-\delta)K$ with $A = 3$, $\alpha = 0.4$ and $\delta = 0.5$; $\pi(\theta = 0) = \pi(\theta = 1) = 1/2$; $G = 0.1$; $k_0 = 1$.

Table 1 presents the relationship between the coefficient of relative risk aversion (σ) and the asset return wedge (R/r) implied by the optimal allocation under the above choice of the parameters' values.

Table 1: Asset Return Wedge

σ	0.25	0.5	0.75	1	1.5	2	5
R/r	1.06	1.10	1.12	1.19	1.22	1.27	1.36

As we observe, the asset wedge increases with the coefficient of relative risk aversion. Intuitively, when the agents become more risk averse, the incentive problem is more severe, which results in an increase of the incentive cost of human capital investment relative to the incentive cost of physical capital investment. This translates into a larger wedge between marginal returns of these assets.

6.2 The Role of Private Information

In this subsection, we study the quantitative differences between our model, the public human capital investment model, and the model with no private information. The three economies we compare have the same preferences, technology, and endowments. The only difference between them is the degree of private information present.

First, we compute the optimal allocation $(c^*, i^*, k_1^*, l^*, K^*, Y^*)$, the intertemporal wedge, the asset return wedge, optimal wealth and labor taxes (τ^*, ϕ^*) and the implied interest and wage rates (r^*, w^*) in our economy with private human capital investment and private skills. Second, we modify our environment by assuming that the agent's human capital investment is public (and thus so is consumption), and then we compute the allocation and the tax code that are optimal in this public human capital investment economy. Finally, we modify our environment by assuming that both the agent's human capital investment and the productivity of this investment are public (and thus so is the agent's skill). Then we compute the allocation and tax code that are optimal in such a full information economy.[8]

We take $\beta = 0.9$; $u(c) = \log(c)$; $v(l) = -(l)^2$; $f(h) = 0.2 + h^\gamma$ with $\gamma = 0.5$; $F(K, Y) = AK^\alpha Y^{1-\alpha} + (1-\delta)K$ with $A = 3$, $\alpha = 0.4$ and $\delta = 0.5$; $\pi(\theta = 0) = \pi(\theta = 1) = 1/2$; $G = 0.1$; $k_0 = 1$.[9] Table 2 presents the optimal allocation, the implied gross interest rate on capital, the wage, and the optimal taxes in these three economies under the above choice of the parameters' values.

[8] In the public human capital investment economy and full information economy, the optimal capital and labor taxes are expressed as a function of the individual component of optimal allocation (c, i, k, l) and prices it implies $r = F_1\left(k, \sum_{\theta\in\Theta} \pi(\theta) f(\theta i)\, l(\theta)\right)$ and $w = F_2\left(k, \sum_{\theta\in\Theta} \pi(\theta) f(\theta i)\, l(\theta)\right)$ as: $\tau(\theta) = 1 - u'(c_0)/r\beta u'(c_1(\theta))$ and $\phi(\theta) = (1-\tau(\theta))rk + f(\theta i)l(1)w - c_1(\theta)$ for $\theta \in \Theta$.

[9] All results presented in this section remain valid across all other parameterizations of our model we considered.

Table 2: Optimal Allocations, Prices, Taxes and Wedges

	Economy with Private Human Capital Investment and Private Skills	*Economy with Public Human Capital Investment and Private Skills*	*Economy with Full Information*
c_0	0.620	0.609	0.609
i	0.051	0.069	0.074
k_1	0.327	0.321	0.316
$c_1(0)$	0.560	0.600	0.759
$c_1(1)$	0.861	0.861	0.759
$l(0)$	0.375	0.324	0.290
$l(1)$	0.589	0.616	0.685
r	1.290	1.333	1.385
w	2.377	2.294	2.203
$\tau(0)$	0.276	0.178	0
$\tau(1)$	-0.276	-0.178	0
$\phi(0)$	-0.075	-0.099	-0.192
$\phi(1)$	0.275	0.299	0.392
$E\left[\frac{\beta u'(c_1)}{u'(c_0)}r\right]$	1.061	1.033	1.000
R/r	1.195	1.008	1.000

We observe that the spread of period-1 consumption, the intertemporal wedge, and the volatility of wealth taxes increase with the degree of private information. In the full information economy, the incentive problem is absent so at the optimum there is no spread of period-1 consumption, no intertemporal wedge, and the marginal optimal wealth taxes are zero. With private information, incentives require a spread in consumption and a positive intertemporal wedge. This wedge increases across environments as more choice variables become private. This translates into a larger and larger spread of the optimal marginal tax rates on capital.

We also observe that more private information leads to less human capital investment. As the incentive problem becomes more severe, the incentive cost of human capital investment increases. Therefore, at the optimum, there is less human capital investment in the economy with more private information.

The asset return wedge is much smaller in the economy with public human capital investment relative to the economy in which human capital investment is private. Intuitively, when human capital investment is private information in addition to human capital, the incentive problem is more severe and the constraint optimal allocation is more inefficient relative to the full insurance (first best) allocation at which the returns on all assets are equal. This relative inefficiency translates into a larger human capital premium.

7 Conclusion

Our analysis of the optimal taxation problem in a Mirrlees economy with endogenous, risky and private human capital shows the following: (1) It is possible to implement the optimum with a capital and labor income tax system in which capital taxes are linear in capital and labor income taxes depend only on labor income. (2) Endogenous skill determination does not affect the average optimal capital tax rate, which continues to be equal to zero. Instead, as we show both numerically and analytically, it increases the volatility of optimal marginal capital tax rates. (3) It is optimal for the return on human capital investment to be strictly larger than the return on physical capital investment.

In this paper, we study a two-period environment. An extension of our model to a general dynamic Mirrlees economy is not straightforward for two reasons. First, in a multiperiod setting, human capital becomes a hidden state variable that evolves stochastically over time, presumably with persistence. In general, optimal allocations are very difficult to characterize in such environments. Second, as our analysis of the two-period model shows, when intertemporal marginal rates of substitution are private, a simple invocation of incentive compatibility of the optimal allocation is not sufficient to ensure the incentive compatibility of the market mechanism in a fiscal decentralization. In particular, some of the off-equilibrium strategies available to agents in the market mechanism provide strictly more utility than their counterparts in the direct revelation mechanism, in which the optimal allocation is determined. A general method for finding the values that these off-equilibrium strategies attain in the market mechanism is not presently available. An extension of our model to an environment with a more complicated temporal and stochastic structure would involve a computation of values of very many off-equilibrium strategies, which would quickly become intractable.

Appendix

Proof of Lemma 1

Necessity If allocation (c, i, k_1, l, K, Y) is IC, then (11) must hold. If it did not, then the strategy of investing i and providing $f(i)l(1)$ units of effective labor irrespectively of θ would yield more utility to the agent than the truthful investment/revelation strategy of investing the recommended amount i and providing $f(\theta)l(\theta)$ units of effective labor in state θ, which would violate the incentive compatibility of (c, i, k_1, l, K, Y).

If allocation (c, i, k_1, l, K, Y) is IC, then (13) must hold. If it did not, the strategy of investing $\widetilde{i}$ and providing $f(0)l(0)$ units of effective labor irrespectively of θ would yield more utility to the agent than the truthful investment/revelation strategy, which would violate the incentive compatibility of (c, i, k_1, l, K, Y).

If allocation (c, i, k_1, l, K, Y) is IC, then (12) must hold. If it did not, then there would be an $i' \neq i$ such that the strategy of investing i' and providing $f(\theta)l(\theta)$ units of effective labor in state θ would yield more utility to the agent than the truthful investment/revelation strategy, and this would violate the incentive compatibility of (c, i, k_1, l, K, Y).

Sufficiency Note that since $w(0, i')$ does not depend on i', we can write function V given in (7) as

$$\beta\pi(1)w(0, i') + \max_{0 \le i' \le c_0 + i} u(c_0 + i - i') + \beta\pi(1)w(1, i').$$

Under (11), we have $w(0, i') = u(c_1(0)) + v(l(0))$. Thus, an allocation (c, i, k_1, l, K, Y) that satisfies (11) is incentive compatible (satisfies the IC constraint (6)) if and only is it satisfies the following condition

$$u(c_0) + \beta\pi(1)\left[u(c_1(1)) + v(l(1))\right] \geq$$
$$\max_{0 \leq i' \leq c_0 + i} u(c_0 + i - i') + \beta\pi(1) \max\left\{ u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i')}\right) ,\, u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i')}\right) \right\}.$$

Exchanging the order of max operations, we get the following equivalent expression of the IC condition:

$$u(c_0) + \beta\pi(1)\left[u(c_1(1)) + v(l(1))\right] \geq$$
$$\max\left\{ \begin{array}{c} \max\limits_{0 \leq i' \leq c_0 + i} u(c_0 + i - i') + \beta\pi(1)\left[u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i')}\right)\right] , \\ \max\limits_{0 \leq i' \leq c_0 + i} u(c_0 + i - i') + \beta\pi(1)\left[u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i')}\right)\right] \end{array} \right\}.$$

This constraint, in turn, is satisfied if and only if the following two constraints are satisfied:

$$u(c_0) + \beta\pi(1)\left[u(c_1(1)) + v(l(1))\right] \geq \tag{43}$$
$$\max_{0 \leq i' \leq c_0 + i} u(c_0 + i - i') + \beta\pi(1)\left[u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i')}\right)\right] ,$$

$$u(c_0) + \beta\pi(1)\left[u(c_1(1)) + v(l(1))\right] \geq \tag{44}$$
$$\max_{0 \leq i' \leq c_0 + i} u(c_0 + i - i') + \beta\pi(1)\left[u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i')}\right)\right] .$$

The expressions under the max operator on the right-hand sides of (43) and (44) are strictly concave in i'. Therefore, the maximum in (43) is attained by i' equal to $\widetilde{i}$ given in (14). Also, by (12), the maximum in (44) is attained by i' equal to i. That the allocation (c, i, k_1, l, K, Y) satisfies (43) is then implied by condition (13). Also (c, i, k_1, l, K, Y) satisfies (44) because the right-hand and the left-hand sides of (44) are identical when i' is equal to i. $\square$

Proof of Lemma 2

We prove this lemma in a series of steps. In Step 1, we show that the constraint (12) does not bind in problem P2. Then, in Step 2, we show that the ex ante IC condition (13) binds in P2. Finally, in Step 3, we show that constraint (11) does not bind in problem P2.

Step 1 We prove that constraint (12) does not bind in problem P2 by considering a relaxed version of problem P2 in which (12) is dropped. We then show that the solution to the relaxed problem satisfies (12).

Consider a relaxed version of problem P2 in which (12) is dropped. Due to the assumed strict convexity of preferences, this problem P2 has a unique solution. Due to the Inada conditions, this solution is interior. The Lagrangian of this relaxed problem is

$$
\begin{aligned}
& u(c_0) + \beta \sum_{\theta \in \Theta} \pi(\theta) \left[u(c_1(\theta)) + v(l(\theta)) \right] \\
& + \alpha \left[u(c_0) - u(c_0 + i - \widetilde{i}) + \pi(1)\beta \left[u(c_1(1)) - u(c_1(0)) + v(l(1)) - v\left(\frac{f(0)l(0)}{f(\widetilde{i})} \right) \right] \right] \\
& + \mu \left[u(c_1(0)) + v(l(0)) - u(c(1)) - v\left(\frac{f(i)l(1)}{f(0)} \right) \right] \\
& + \lambda_0 \left[k_0 - c_0 - i - k_1 \right] \\
& + \lambda_1 \left[F\left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right) - \sum_{\theta} \pi(\theta) c_1(\theta) - G \right]
\end{aligned}
$$

where $\widetilde{i}$ is a function of $c_0, i, l(0)$ given implicitly by

$$
u'(c_0 + i - \widetilde{i}) = -\pi(1)\beta v'\left(\frac{f(0)l(0)}{f(\widetilde{i})} \right) \frac{f(0)l(0)}{f(\widetilde{i})^2} f'(\widetilde{i}) \tag{45}
$$

and where α, μ, λ_0, and λ_1 are the Lagrange multipliers of the respective constraints. Taking the FO conditions of this Lagrangian with respect to $c_0, i, l(0)$, by the Envelope Theorem, the terms that involve the derivative $d\widetilde{i}/dx$ drop out. The FO conditions for the unique interior solution to the relaxed planner problem P2, then, are as follows:

$$
[c_0 :]\ u'(c_0) + \alpha \left[u'(c_0) - u'(c_0 + i - \widetilde{i}) \right] - \lambda_0 = 0, \tag{46}
$$

$$
[c_1(0) :] \quad \pi(0)\beta u'(c_1(0)) - \alpha\beta\pi(1)u'(c_1(0)) + \mu u'(c_1(0)) - \lambda_1 \pi(0) = 0, \tag{47}
$$

$$
[c_1(1) :] \quad \pi(1)\beta u'(c_1(1)) + \alpha\beta\pi(1)u'(c_1(1)) - \mu u'(c_1(1)) - \lambda_1 \pi(1) = 0, \tag{48}
$$

$$
\begin{aligned}
[i :] \quad & -\alpha u'(c_0 + i - \widetilde{i}) - \mu v'\left(\frac{f(i)l(1)}{f(0)} \right) \frac{f'(i)l(1)}{f(0)} \\
& -\lambda_0 + \lambda_1 \pi(1) f'(i) l(1) F_2 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right) = 0,
\end{aligned} \tag{49}
$$

$$
[k_1 :] \quad -\lambda_0 + \lambda_1 F_1 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right) = 0, \tag{50}
$$

$$
\begin{aligned}
[l(0) :] \quad & \pi(0)\beta v'(l(0)) - \alpha\pi(1)\beta v'\left(\frac{f(0)l(0)}{f(\widetilde{i})} \right) \frac{f(0)}{f(\widetilde{i})} + \mu v'(l(0)) \\
& + \lambda_1 \pi(0) f(0)\, F_2 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right) = 0,
\end{aligned} \tag{51}
$$

$$
\begin{aligned}
[l(1) :] \quad & \pi(1)\beta v'(l(1)) + \alpha\pi(1)\beta v'(l(1)) \\
-\mu v'\left(\frac{f(i)l(1)}{f(0)} \right) \frac{f(i)}{f(0)} & + \pi(1) f(i) \lambda_1 F_2 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right) = 0.
\end{aligned} \tag{52}
$$

In order to show that (12) does not bind in the unrelaxed planner's problem P1, it is sufficient to show that the solution to the relaxed problem P2 satisfies (12). Indeed, combining (46) with (49) yields

$$\frac{(1+\alpha)u'(c_0)}{f'(i)l(1)} = \lambda_1 \pi(1) F_2 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right) - \mu v' \left(\frac{f(i)l(1)}{f(0)} \right) \frac{1}{f(0)}.$$

The FO condition (52) implies

$$-\frac{\pi(1)\beta v'(l(1))(1+\alpha)}{f(i)} = \lambda_1 \pi(1) F_2 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right) - \mu v' \left(\frac{f(i)l(1)}{f(0)} \right) \frac{1}{f(0)}.$$

Combining the above two equations yields

$$u'(c_0) = -\pi(1)\beta \frac{v'(l(1))}{f(i)} f'(i) l(1),$$

which is condition (12). Hence, condition (12) does not bind at the optimum.

Step 2 We prove by contradiction that, in problem P2, the ex ante IC condition (13) binds. If (13) does not bind in P2, then the solution to a relaxed problem P2 in which (13) is dropped must satisfy (13). We show that if (13) is dropped from P2, then, taking into account that (12) does not bind, the full insurance allocation (sometimes also called the first best allocation) solves that problem. But full insurance violates (13), hence (13) must bind.

From Step 1, we know that the optimal allocation is a unique interior solution of the relaxed problem P2 which is formed from P1 by dropping the constraint (12). Suppose that at the constrained optimal allocation $(c^*, k_1^*, i^*, l^*, K^*, Y^*)$, the ex-ante IC constraint (13) is not binding.

Consider the full insurance allocation (sometimes also called the first best allocation), which is the unique resource feasible allocation that satisfies $c_1(0) = c_1(1)$ (full insurance of consumption) and

$$u'(c_0) = \beta u'(c_1(1)) \pi(1) f'(i) l(1) F_2 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right), \tag{53}$$

$$u'(c_0) = \beta u'(c_1(1)) F_1 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right), \tag{54}$$

$$-v'(l(0)) = u'(c_1(1)) f(0)\, F_2 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right), \tag{55}$$

$$-v'(l(1)) = u'(c_1(1)) f(i) F_2 \left(k_1, \sum_{\theta \in \Theta} \pi(\theta) f(\theta i)\, l(\theta) \right). \tag{56}$$

Denote the full insurance allocation by $(c^{fi}, k_1^{fi}, i^{fi}, l^{fi}, K^{fi}, Y^{fi})$. From (53) through (56) we get that

$$l^{fi}(1) > l^{fi}(0), \quad i^{fi} > 0. \tag{57}$$

If the incentive compatibility condition (11) does not bind at the solution to P2, i.e., if its Lagrange multiplier $\alpha = 0$, we check directly that the full insurance allocation satisfies the FO conditions of P2, (46) through (52), with $\mu = 0$, $\lambda_0 = u'(c_0^{fi})$, and $\lambda_1 = \beta u'(c_1^{fi}(1)) = \beta u'(c_1^{fi}(0))$. This, combined with the fact that the full insurance allocation is resource feasible, implies that the full insurance allocation is the unique solution to the planner's problem P2 when the constraint (13) is not binding. However, this yields a contradiction as the the full insurance allocation does not satisfy the constraint (13). To see this, note that the fact that $\widetilde{i}$ maximizes the right-hand side of (43) implies that

$$u(c_0 + i - \widetilde{i}) + \beta\pi(1)\left[u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(\widetilde{i})}\right)\right] \geq u(c_0 + i) + \beta\pi(1)\left[u(c_1(0)) + v\left(l(0)\right)\right].$$

Therefore, for an allocation to satisfy (13), it is necessary that constraint (13) implies that at the incentive compatible allocation

$$u(c_0) + \beta\pi(1)\left[u(c_1(1)) + v(l(1))\right] \geq u(c_0 + i) + \beta\pi(1)\left[u(c_1(0)) + v\left(l(0)\right)\right]. \tag{58}$$

But, using (57), we get that

$$u(c_0^{fi}) + \beta\pi(1)\left[u(c_1^{fi}(1)) + v(l^{fi}(1))\right] < u(c_0^{fi} + i^{fi}) + \beta\pi(1)\left[u(c_1^{fi}(0)) + v\left(l^{fi}(0)\right)\right],$$

which implies that the full insurance allocation violates (58), and thus it also violates the IC constraint (13). Therefore, we conclude, the ex ante IC constraint (13) must bind at the solution to P2.

Step 3 We prove that constraint (11) does not bind in problem P2 as follows. First, in Step 3a, we show that any incentive feasible allocation that satisfies both (11) and (13) as equalities is a so-called bunching allocation at which both ex post types of agents provide the same amount of effective labor and consume equal amounts of the consumption good. Then, in Step 3b, we show that the bunching allocation does not solve P2. This means that at least one of these two constraints, (11) and (13), is slack at the solution to P2. By Step 2, (13) binds. Thus, constraint (11) must be slack at the optimum.

Step 3a If (13) and (11) both bind in problem P2, then, by complementary slackness, they must both be satisfied as equalities. We first show that any incentive feasible allocation that satisfies both (13) and (11) as equalities is a so-called bunching allocation. Then, we show that a bunching allocation can never be optimal.

If (11) holds as equality, then, as u is strictly increasing and v is strictly decreasing, three cases are possible:

1) $c_1(0) > c_1(1)$ and $f(0)l(0) > f(i)l(1)$,

2) $c_1(0) < c_1(1)$ and $f(0)l(0) < f(i)l(1)$,

3) $c_1(0) = c_1(1)$ and $f(0)l(0) = f(i)l(1)$.

Case 1. We have that (11) holding as equality, $c_1(0) > c_1(1)$, and $f(0)l(0) > f(i)l(1)$ imply that for any $i' > 0$

$$u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i')}\right) > u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i')}\right).$$

In particular, for $i' = i$ we get

$$u(c_0) + \beta\pi(1)\left[u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i)}\right)\right] > u(c_0) + \beta\pi(1)\left[u(c_1(1)) + v(l(1))\right].$$

Because $\widetilde{i}$ satisfies (14), i.e., maximizes the value of the deviation strategy of announcing $\theta = 0$ when $\theta = 1$, we have

$$u(c_0 + i - \widetilde{i}) + \beta\pi(1)\left[u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(\widetilde{i})}\right)\right] \geq u(c_0) + \beta\pi(1)\left[u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i)}\right)\right].$$

The last two inequalities put together contradict (13).

Case 2. We have that (11) holding as equality, $c_1(0) < c_1(1)$, and $f(0)l(0) < f(i)l(1)$ imply that for any $i' > 0$

$$u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i')}\right) < u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i')}\right).$$

In particular, for $i' = \widetilde{i}$ we get

$$u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(\widetilde{i})}\right) < u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(\widetilde{i})}\right).$$

Combining the above strict inequality with (13) holding with equality and subtracting $\beta\pi(1)u(c_1(1))$ from both sides, we get

$$u(c_0) + \beta\pi(1)v\left(l(1)\right) < u(c_0 + i - \widetilde{i}) + \beta\pi(1)v\left(\frac{f(i)l(1)}{f(\widetilde{i})}\right),$$

which contradicts the assumption that i satisfies (12).

Case 3. If $f(0)l(0) = f(i)l(1)$, then (12) and (14) imply that $\widetilde{i} = i$. Condition (13) reduces then to

$$u(c_1(1)) + v\left(\frac{f(i)l(1)}{f(i)}\right) \geq u(c_1(0)) + v\left(\frac{f(0)l(0)}{f(i)}\right).$$

Both this condition and (11) are (trivially) satisfied with equality because $c_1(0) = c_1(1)$ and $f(0)l(0) = f(i)l(1)$.

We have thus shown that if an allocation satisfies both (13) and (11) as equalities, then $c_1(0) = c_1(1)$ and $f(0)l(0) = f(i)l(1)$. We will refer to any such allocation as a bunching allocation and denote it by $(c^b, i^b, k_1^b, l^b, K^b, Y^b)$.

Step 3b Now we show that no bunching allocation is optimal. If a bunching allocation is not resource feasible, we are done. Suppose then it is RF. And suppose it is optimal. Then, there must exist Lagrange multipliers α, μ, λ_0, λ_1, such that the bunching allocation $(c^b, i^b, k_1^b, l^b, K^b, Y^b)$ satisfies the FO conditions (46) through (52). In addition, under the assumption that (13) and (11) bind, all those Lagrange multipliers must be strictly positive.

The FO conditions (47) and (48) imply that

$$\begin{aligned} u'(c_1^b(0)) &= \frac{\lambda_1}{\beta\left(1 - \alpha\frac{\pi(1)}{\pi(0)} + \frac{\mu}{\pi(0)\beta}\right)}, \\ u'(c_1^b(1)) &= \frac{\lambda_1}{\beta\left(1 + \alpha - \frac{\mu}{\pi(1)\beta}\right)}, \end{aligned}$$

which, given the strict concavity of u and the fact that $c_1^b(0) = c_1^b(1)$, implies that

$$\frac{\mu}{\pi(1)\beta} = \alpha. \tag{59}$$

Using the fact that $f(i^b)l^b(1) = f(0)l^b(0)$, $i^b = \widetilde{i} > 0$, and (59), we can write the FO conditions (51) and (52) as

$$\lambda_1 F_2\left(k_1, \sum_{\theta\in\Theta} \pi(\theta) f(\theta i)\, l(\theta)\right) = \frac{\mu v'(l(1))}{\pi(0)f(i)} - \frac{\mu v'(l(0))}{\pi(0)f(0)} - \frac{\beta v'(l(0))}{f(0)}, \tag{60}$$

$$\frac{\beta v'(l(1))}{f(i)} + \frac{\mu v'(l(1))}{\pi(1)f(i)} - \frac{\mu v'(l(0))}{\pi(1)f(0)} + \lambda_1 F_2\left(k_1, \sum_{\theta\in\Theta} \pi(\theta) f(\theta i)\, l(\theta)\right) = 0. \tag{61}$$

Substituting (60) into (61) and cancelling out terms, we get the following condition:

$$\frac{v'\left(\frac{f(0)l^b(0)}{f(i^b)}\right)}{f(i^b)} = \frac{v'\left(\frac{f(0)l^b(0)}{f(0)}\right)}{f(0)},$$

which, as $v'\left(\frac{y}{f(i)}\right)/f(i)$ is strictly increasing in i, is a contradiction.

We have shown in Step 3b that no bunching allocation $(c^b, i^b, k_1^b, l^b, K^b, Y^b)$ is constrained optimal. We have also shown in Step 3a that there is no non-bunching allocation that satisfies both (11) and (13) as equalities. Therefore, it is not true that at the solution to P2 both IC constraints (11) and (13) bind. By Step 2, (13) does bind. Thus, (11) does not. □

Proof of Proposition 1

To obtain (21), we combine the FO conditions (46)-(52) of the planner's problem to get

$$\frac{u'(c_0^*)}{\beta u'(c_1^*(0))} = \left(\frac{1 - \alpha\frac{\pi(1)}{\pi(0)}}{1+\alpha}\right) R^*, \tag{62}$$

$$\frac{u'(c_0^*)}{\beta u'(c_1^*(1))} = R^*, \tag{63}$$

$$\lambda_1 r^* = \lambda_1 R^* - \alpha u'(c_0^* + i^* - \widetilde{i}^*), \tag{64}$$

$$\lambda_0 = \lambda_1 r^*. \tag{65}$$

Combining (62) and (63) we get

$$\sum_{\theta\in\Theta} \pi(\theta) \frac{u'(c_0^*)}{\beta u'(c_1^*(\theta))} = \frac{R^*}{1+\alpha}. \tag{66}$$

The first order conditions (46) and (49) of the planner's problem imply that

$$\frac{R^*}{1+\alpha} = \frac{u'(c_0^*)}{\lambda_1}. \tag{67}$$

Substituting this into (66) yields

$$\sum_{\theta\in\Theta}\frac{\pi(\theta)}{\beta u'(c_1^*(\theta))}=\frac{1}{\lambda_1}. \tag{68}$$

The first order conditions (46) and (65) imply that

$$\lambda_1=\frac{u'(c_0)+\alpha\left[u'(c_0)-u'(c_0+i-\widetilde{i})\right]}{r^*}.$$

Substituting this into (68) and rearranging yields the modified Rogerson equation (21)

$$u'(c_0)+\alpha\left[u'(c_0)-u'(c_0+i-\widetilde{i})\right]=\frac{\beta r^*}{E[1/u'(c_1^*)]}.$$

In order to confirm the formula for α given in (22), we use (63) to eliminate R^* from (62) and solve for α, which yields

$$\alpha=\frac{\frac{u'(c_1^*(0))}{u'(c_1^*(1))}-1}{1+\frac{\pi(1)}{\pi(0)}\frac{u'(c_1^*(0))}{u'(c_1^*(1))}}.$$

□

Proof of Proposition 2

Condition (31) follows directly from (66). □

Proof of Proposition 3

In order to show that the tax system (τ^*,ϕ^*) is optimal, we need to show that if taxes are (τ^*,ϕ^*), then the optimal allocation $(c^*,k_1^*,i^*,l^*,K^*,Y^*)$ is an equilibrium allocation.
We claim that it indeed is under equilibrium prices r and w given by the optimal marginal returns r^* and w^*, respectively.

First, as the optimal allocation $(c^*,k_1^*,i^*,l^*,K^*,Y^*)$ is resource feasible, the market clearing conditions are satisfied. Also, the prices (r^*,w^*) satisfy the competitive pricing condition of the equilibrium definition. Thus, all that remains to be checked is that, given the tax system (τ^*,ϕ^*) and prices (r^*,w^*), the individual component of the optimal allocation, (c^*,k_1^*,i^*,l^*), is a solution to the problem of agents' utility maximization .

Turning to this problem, then, we note that under simplified taxes (τ^*,ϕ^*) there are four strategies of effective labor supply that do not lead to zero consumption. These are

1. $s=(0,1)$, which corresponds to providing low effective labor $f(0)l^*(0)$ when $\theta=0$ and high effective labor $f(i^*)l^*(1)$ when $\theta=1$.

2. $s=(0,0)$, which corresponds to providing low effective labor $f(0)l^*(0)$ independently of the value of θ.

3. $s=(1,1)$, which corresponds to providing high effective labor $f(i^*)l^*(1)$ independently of the value of θ.

4. $s=(1,0)$, which corresponds to providing high effective labor $f(i^*)l^*(1)$ when $\theta=0$ and low effective labor $f(0)l^*(0)$ when $\theta=0$.

As any other effective labor supply strategy leads to zero consumption, the optimal effective labor supply plan of any agent will be to follow one of the above four strategies. Thus, the overall solution to the utility maximization problem under prices (r^*, w^*) and taxes (τ^*, ϕ^*) can be obtained by first optimizing with respect to variables $c_0, c_1(\theta), k_1, i$ conditional on the effective labor supply strategy $n = 1, 2, 3, 4$ and then comparing the values attained by each of the four effective labor supply strategies.

Below we show that, under prices (r^*, w^*) and taxes (τ^*, ϕ^*), strategy 1 yields as much utility as any of the other three strategies and thus is weakly optimal in the agent's equilibrium problem. Also, agents' individually optimal choices of $c_0, c_1(\theta), k_1, i$ under strategy 1 coincide with the socially optimal levels $c^*, c_1^*(\theta), k_1^*, i^*$ and the levels of the labor effort provided by agents under this strategy 1 coincide with the optimal levels $l^*(\theta)$ for both θ.

Strategy 1. Suppose that an agent decides to provide low effective labor $f(0)l^*(0)$ when $\theta = 0$ and high effective labor $f(i^*)l^*(1)$ when his $\theta = 1$. Conditional on these effective labor supply choices, at prices (r^*, w^*) and taxes (τ^*, ϕ^*), the utility maximization problem is given as follows:

$$\max_{c_0, c_1(\theta), k_1, i \geq 0} u(c_0) + \beta\pi(0)[u(c_1(0)) + v(l^*(0))] + \beta\pi(1)\left[u(c_1(1)) + v\left(\frac{f(i^*)l^*(1)}{f(i)}\right)\right]$$

subject to

$$\begin{aligned} c_0 + i + k_1 &\leq k_0, \\ c_1(0) &\leq (1 - \tau^*(0))r^* k_1 + w^* f(0) l^*(0) - \phi^*(0), \\ c_1(1) &\leq (1 - \tau^*(1))r^* k_1 + w^* f(i^*) l^*(1) - \phi^*(1). \end{aligned}$$

The solution of the above problem is characterized by the following system of equations

$$\begin{aligned} u'(c_0) &= \pi(0)(1 - \tau^*(0))\beta r^* u'(c_1(0)) + \pi(1)(1 - \tau^*(1))\beta r^* u'(c_1(1)), \\ u'(c_0) &= -\beta\pi(1)v'\left(\frac{f(i^*)l^*(1)}{f(i)}\right)\frac{f(i^*)l^*(1)}{f(i)^2}f'(i), \\ c_0 + i + k_1 &= k_0, \\ c_1(0) &= (1 - \tau^*(0))r^* k_1 + w^* f(0) l^*(0) - \phi^*(0), \\ c_1(1) &= (1 - \tau^*(1))r^* k_1 + w^* f(i^*) l^*(1) - \phi^*(1), \end{aligned}$$

where the first two equations are FO conditions with respect to k_1 and i. Substituting the expressions in (37), (38) and (39) for simplified taxes yields

$$\begin{aligned} u'(c_0) &= \frac{\pi(0)u'(c_0^* + i^* - \widetilde{i}^*)}{u'(c_1^*(0))}u'(c_1(0)) + \frac{u'(c_0^*) - \pi(0)u'(c_0^* + i^* - \widetilde{i}^*)}{u'(c_1^*(1))}u'(c_1(1)), \\ u'(c_0) &= -\beta\pi(1)v'\left(\frac{f(i^*)l^*(1)}{f(i)}\right)\frac{f(i^*)l^*(1)}{f(i)^2}f'(i), \\ c_0 + i + k_1 &= k_0, \\ c_1(0) &= c_1^*(0) + (1 - \tau^*(0))r^*(k_1 - k_1^*), \\ c_1(1) &= c_1^*(1) + (1 - \tau^*(1))r^*(k_1 - k_1^*). \end{aligned}$$

We check directly that the optimal consumption levels c_0^*, $c_1^*(0)$, $c_1^*(1)$ satisfy the first FO condition. That c_0^* and i^* solve the second FO condition follows from the fact that the optimal allocation satisfies (12). Since the optimal allocation is resource feasible, it satisfies the ex ante budget constraint. Finally, we check directly that $c_1^*(0)$, $c_1^*(1)$ and k_1^* satisfy the ex post budget constraints for both θ. By strict concavity of the objective, $(c_0^*, c_1^*(0), c_1^*(1), i^*, k_1^*)$ is the unique solution to the utility maximization problem under effective labor supply strategy 1, prices (r^*, w^*) and taxes (τ^*, ϕ^*). Using the identity $f(\theta i^*)l^*(\theta)/f(\theta i) = l^*(\theta)$, we get that the maximal utility attained under strategy 1 is equal to

$$u(c_0^*) + \beta \sum_{\theta \in \Theta} \left[\pi(\theta) u(c_1^*(\theta)) + v(l^*(\theta)) \right].$$

We have thus shown that, under taxes (τ^*, ϕ^*) and prices (r^*, w^*), the individually optimal choices of investment, consumption and labor coincide with the individual component of the optimal allocation $(c_0^*, c_1^*(0), c_1^*(1), i^*, k_1^*)$. All that remains to be shown to complete the proof is that strategy 1 weakly dominates strategies 2, 3, and 4 in agents' equilibrium utility maximization problem.

Strategy 2. Suppose that an agent decides to provide low effective labor $f(0)l^*(0)$ independently of the value of his θ. Conditional on these effective labor supply choices, at prices (r^*, w^*) and taxes (τ^*, ϕ^*), the utility maximization problem is given as follows:

$$\max_{c_0, c_1(\theta), k_1, i \geq 0} u(c_0) + \beta\pi(0)[u(c_1(0)) + v(l^*(0))] + \beta\pi(1) \left[u(c_1(1)) + v\left(\frac{f(0)l^*(0)}{f(i)} \right) \right]$$

subject to

$$\begin{aligned} c_0 + i + k_1 &\leq k_0, \\ c_1(0) &\leq (1 - \tau^*(0)) r^* k_1 + w^* f(0) l^*(0) - \phi^*(0), \\ c_1(1) &\leq (1 - \tau^*(0)) r^* k_1 + w^* f(0) l^*(0) - \phi^*(0). \end{aligned}$$

Let the solution to this problem be denoted by $(c_0^2, c_1^2(\theta), i^2, k_1^2)$. At the solution, the implied labor levels are $l^2(0) = l^*(0)$ and $l^2(1) = \frac{f(0)l^*(0)}{f(i^2)}$. The FO conditions and budget equations are sufficient. Taking FO conditions and using the fact the budget constraints bind, we get that $(c_0^2, c_1^2(\theta), i^2, k_1^2)$ solves the following system of equations:

$$\begin{aligned} u'(c_0) &= \pi(0)(1 - \tau^*(0))\beta r^* u'(c_1(0)) + \pi(1)(1 - \tau^*(0))\beta r^* u'(c_1(1)), \\ u'(c_0) &= -\beta\pi(1) v' \left(\frac{f(i^*)l^*(1)}{f(i)} \right) \frac{f(i^*)l^*(1)}{f(i)^2} f'(i), \\ c_0 + i + k_1 &= k_0, \\ c_1(0) &= (1 - \tau^*(0)) r^* k_1 + w^* f(0) l^*(0) - \phi^*(0), \\ c_1(1) &= (1 - \tau^*(0)) r^* k_1 + w^* f(0) l^*(0) - \phi^*(0). \end{aligned}$$

Using (37) and (39), we eliminate taxes $\tau^*(0)$ and $\phi^*(0)$ from these conditions and get the following set of sufficient conditions:

$$
\begin{aligned}
u'(c_0) = \pi(0)\frac{u'(c_0^* + i^* - \widetilde{i}^*)}{u'(c_1^*(0))}u'(c_1(0)) + \pi(1)\frac{u'(c_0^* + i^* - \widetilde{i}^*)}{u'(c_1^*(0))}u'(c_1(1)),\\
u'(c_0) = -\beta\pi(1)v'\left(\frac{f(0)l^*(0)}{f(i)}\right)\frac{f(0)l^*(0)}{f(i)^2}f'(i),\\
c_0 + i + k_1 = k_0,\\
c_1(0) = c_1^*(0) + \frac{u'(c_0^* + i^* - \widetilde{i}^*)}{\beta u'(c_1^*(0))}(k_1 - k_1^*),\\
c_1(1) = c_1^*(0) + \frac{u'(c_0^* + i^* - \widetilde{i}^*)}{\beta u'(c_1^*(0))}(k_1 - k_1^*).
\end{aligned}
$$

We claim that $(c_0^2, c_1^2(0), c_1^2(1), i^2, k_1^2) = (c_0^* + i^* - \widetilde{i}^*, c_1^*(0), c_1^*(0), \widetilde{i}^*, k_1^*)$ with the associated choice of labor given by $(l^2(0), l^2(1)) = \left(l^*(0), \frac{f(0)l^*(0)}{f(\widetilde{i}^*)}\right)$ is the solution to the utility maximization problem under strategy 2, i.e., that these values of c, i, and k_1 satisfy the above sufficient conditions. That $c_0^2 = c_0^* + i^* - \widetilde{i}^*$ and $i^2 = \widetilde{i}^*$ solve the second FO condition follows from the fact that the optimal allocation satisfies (14). We check directly that the remaining four conditions are satisfied, too. Thus, the maximal utility attained by strategy 2 is equal to

$$
u(c_0^* + i^* - \widetilde{i}^*) + \beta\left[u(c_1^*(0)) + \pi(0)v(l^*(0)) + \pi(1)v\left(\frac{f(0)l^*(0)}{f(\widetilde{i}^*)}\right)\right].
$$

By Lemma 3, we have that

$$
\begin{aligned}
u(c_0^* + i^* - \widetilde{i}^*) + \beta u(c_1^*(0)) + \beta\left[\pi(0)v(l^*(0)) + \pi(1)v\left(\frac{f(0)l^*(0)}{f(\widetilde{i}^*)}\right)\right] =\\
u(c_0^*) + \beta\sum_{\theta\in\Theta}\left[\pi(\theta)u(c_1^*(\theta)) + v(l^*(\theta))\right],
\end{aligned}
$$

which implies that the strategies 1 and 2 yield the same ex ante utility to an agent, and hence strategy 1 weakly dominates strategy 2.

Strategy 3. Suppose that an agent decides to follow strategy 3 of providing high effective labor $f(i^*)l^*(1)$ independently of the value of his θ. Under this strategy, the optimal consumption, investment and labor decisions of the agent at prices (r^*, w^*) and taxes (τ^*, ϕ^*), denoted (c^3, k_1^3, i^3, l^3), are given as a solution to

$$
\max_{c_0, c_1(\theta), k_1, i\geq 0} u(c_0) + \beta\pi(\theta)\left[u(c_1(\theta)) + v\left(\frac{f(i^*)l^*(1)}{f(\theta i)}\right)\right] \tag{69}
$$

subject to

$$
\begin{aligned}
c_0 + i + k_1 \leq k_0,\\
c_1(0) \leq (1 - \tau^*(1))r^*k_1 + w^*f(i^*)l^*(1) - \phi^*(1),\\
c_1(1) \leq (1 - \tau^*(1))r^*k_1 + w^*f(i^*)l^*(1) - \phi^*(1),
\end{aligned}
$$

with $l^3(0) = \frac{f(i^*)l^*(1)}{f(0)}$ and $l^3(1) = \frac{f(i^*)l^*(1)}{f(i^3)}$. Substituting the budget constraints to the objective, we get the following expression of this maximization problem

$$\max_{k_1, i\geq 0} u(k_0 - k_1 - i) + \beta \left[u(c_1^*(1) + (1-\tau^*(1))r^*(k_1 - k_1^*)) + \pi(0)v\left(\frac{f(i^*)l^*(1)}{f(0)}\right) + \pi(1)v\left(\frac{f(i^*)l^*(1)}{f(i)}\right)\right].$$

We need to show that the value of this problem does not exceed the maximal utility attained by strategy 1. Our argument proceeds in three steps. In the first step, we show directly that strategy 1 dominates strategy 3 in all economies in which the productivity of agents who do not invest in human capital, $f(0)$, is small. In the second step, we show that the values of optimal consumption, investment and labor choices of agents under strategies 1 and 3 converge to a common limit when $f(0)$ tends toward plus infinity. Finally, in the third step, we derive a necessary condition for the strict dominance of strategy 3 over strategy 1 and show, using steps 1 and 2, that this condition is violated in all economies that we study in this paper, i.e., for all values of $f(0)$.

Fix all parameters of the economic environment except of $f(0)$. More precisely, fix u, v, F, π, k_0, G and $f - f(0)$, i.e., the shape of the function f up to a vertical shift. Now $f(0)$ is a single parameter indexing the economies in the subset we obtain by fixing the other parameters. As nothing in our argument depends on a particular way those other parameters are fixed, the argument applies to all such subsets and therefore to all economies we study in this paper.

Step 1 It follows from the Theorem of the Maximum that the optimal allocation and the utility value attained in the planner's problem are continuous with respect to $f(0)$. As prices (r^*, w^*) and taxes (τ^*, ϕ^*) are given as continuous functions of $f(0)$, they also are continuous with respect to $f(0)$. Again from the Maximum Theorem it follows that the utility value of strategy 3 and the maximizers (c^3, k_1^3, i^3, l^3) that attain it are all continuous with respect to $f(0)$. (The same is true for strategies 2 and 4.)

Take $f(0) = 0$. By the Inada conditions, the optimal allocation features a positive effective labor supply at the ex post date. Because of $f(0) = 0$, we have that $f(0)l^*(0) = 0$. Thus, $f(i^*)l^*(1) > 0$. In the utility maximization problem under strategy 3, (69), the term $\beta\pi(0)v\left(\frac{f(i^*)l^*(1)}{f(0)}\right)$ does not depend on any of the choice variables, so it can be taken outside of the maximization operator. The value of the terms that remain inside, at the maximum, is a finite number. However, the value of $\beta\pi(0)v\left(\frac{f(i^*)l^*(1)}{f(0)}\right)$ is negative infinity. Therefore, the value of strategy 3 is negative infinity if $f(0) = 0$. Intuitively, it is infinitely costly for an agent whose productivity is zero to produce a strictly positive amount of output. By continuity, strategy 3 is dominated by strategy 1 for all sufficiently small values of $f(0)$.

Step 2 Using the FO conditions (46) - (52), it is easy to show that as $f(0)$ diverges to plus infinity i^* converges to zero. Intuitively, $f(0)$ becoming large is equivalent to having a lot of human capital at the ex post date. As $f(0)$ increases, the value of an additional unit of human capital ex post decreases relative to the investment opportunity cost ex ante and i^* tends toward zero in the limit. Thus, the incentive problem vanishes as $f(0)$ diverges to plus infinity and the optimal allocation converges to a full insurance allocation. Thus, $|c_1^*(1) - c_1^*(0)| \to 0$, $|l^*(1) - l^*(0)| \to 0$, and, because $i^* > \widetilde{i^*}$, $\widetilde{i^*} \to 0$. Directly from formulas for optimal taxes in (37) - (39) it follows then that capital and labor taxes converge to zero. Therefore, as taxes and effective labor assignments for the two ex post types of agents converge, the investment, consumption and labor choices that are optimal in the agent's problem under effective labor strategy 3 converge to a common limit, the optimal choices under strategy 1. In particular, the difference between the optimal physical capital investments under strategies 3 and 1, k_1^3 and k_1^*, respectively, converges to zero as $f(0)$ tends toward plus infinity: $|k_1^3 - k_1^*| \to 0$.

Step 3 In this last step, we show that strategy 1 weakly dominates strategy 3 in all economies. We derive a necessary condition for a strict domination of strategy 3 over strategy 1 and show that this condition is violated in all economies.

From our analysis of the optimal behavior of an agent who follows effective labor strategy 1, we know that under this strategy consumption c^*, investment i^*, k_1^*, and labor l^* are a unique utility-maximizer. Therefore, investment decisions i^3, k_1^3, and consumption and labor that follow are strictly suboptimal under the effective labor strategy 1. Thus, we have

$$
\begin{aligned}
&u(c_0^*) + \beta\left[\pi(0)\left(u(c_1^*(0)) + v\left(\frac{f(0)l^*(0)}{f(0)}\right)\right) + \pi(1)\left(u(c_1^*(1)) + v\left(\frac{f(i^*)l^*(1)}{f(i^*)}\right)\right)\right] \\
&> u(k_0 - k_1^3 - i^3) + \beta\pi(0)\left(u(c_1^*(0) + (1-\tau^*(0))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(0)l^*(0)}{f(0)}\right)\right) \\
&\quad + \beta\pi(1)\left(u(c_1^*(1) + (1-\tau^*(1))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(i^*)l^*(1)}{f(i^3)}\right)\right).
\end{aligned}
$$

Also, if strategy 3 strictly dominates strategy 1, we have

$$
\begin{aligned}
&u(k_0 - k_1^3 - i^3) + \beta\pi(0)\left(u(c_1^*(1) + (1-\tau^*(1))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(i^*)l^*(1)}{f(0)}\right)\right) \\
&\quad + \beta\pi(1)\left(u(c_1^*(1) + (1-\tau^*(1))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(i^*)l^*(1)}{f(i^3)}\right)\right) \\
&> u(c_0^*) + \beta\left[\pi(0)\left(u(c_1^*(0)) + v\left(\frac{f(0)l^*(0)}{f(0)}\right)\right) + \pi(1)\left(u(c_1^*(1)) + v\left(\frac{f(i^*)l^*(1)}{f(i^*)}\right)\right)\right].
\end{aligned}
$$

Combining the these two inequalities, we get

$$
\begin{aligned}
&u(k_0 - k_1^3 - i^3) + \beta\pi(0)\left(u(c_1^*(1) + (1-\tau^*(1))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(i^*)l^*(1)}{f(0)}\right)\right) \\
&\quad + \beta\pi(1)\left(u(c_1^*(1) + (1-\tau^*(1))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(i^*)l^*(1)}{f(i^3)}\right)\right) \\
&> u(k_0 - k_1^3 - i^3) + \beta\pi(0)\left(u(c_1^*(0) + (1-\tau^*(0))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(0)l^*(0)}{f(0)}\right)\right) \\
&\quad + \beta\pi(1)\left(u(c_1^*(1) + (1-\tau^*(1))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(i^*)l^*(1)}{f(i^3)}\right)\right)
\end{aligned}
$$

which simplifies to

$$
\begin{aligned}
&u(c_1^*(1) + (1-\tau^*(1))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(i^*)l^*(1)}{f(0)}\right) \\
&> u(c_1^*(0) + (1-\tau^*(0))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(0)l^*(0)}{f(0)}\right).
\end{aligned}
\tag{70}
$$

Condition (70) is necessary for the strict dominance of strategy 3 over strategy 1. We argue now that (70) is violated in all economies we study in this paper. I.e., we define the following function

$$\begin{aligned} g(f(0)) &= u(c_1^*(0) + (1-\tau^*(0))r^*(k_1^3 - k_1^*)) + v\left(\frac{f(0)l^*(0)}{f(0)}\right) \\ &\quad -u(c_1^*(1) + (1-\tau^*(1))r^*(k_1^3 - k_1^*)) - v\left(\frac{f(i^*)l^*(1)}{f(0)}\right) \end{aligned}$$

and show that $g(f(0)) \geq 0$ for all $f(0)$.

By Step 1 above, $g(f(0)) > 0$ for all sufficiently small values of $f(0)$. By Step 2 above, $\lim_{f(0)\to\infty} g(f(0)) = 0$. By continuity of g, we get the desired conclusion that $g \geq 0$ if we show that g is monotonically decreasing. The derivative $dg/df(0)$ involves direct and indirect terms that influence g through the values of the elements of the optimal allocation and k_1^3. By Step 2, as $f(0)$ increases, $c_1^*(0)$ tends toward $c_1^*(1)$, $y^*(0) = f(0)l^*(0)$ tends toward $y^*(1) = f(i^*)l^*(1)$ and both $(1-\tau^*(0))r^*(k_1^3 - k_1^*)$ and $(1-\tau^*(1))r^*(k_1^3 - k_1^*)$ tend toward zero. The indirect terms in the derivative $dg/df(0)$ therefore tend to cancel each other out. Thus, we assume that the indirect effects are second order and approximate the value of the derivative $dg/df(0)$ by the value of the direct terms as follows:[10]

$$\frac{dg(f(0))}{df(0)} \approx v'\left(\frac{y^*(0)}{f(0)}\right)\frac{-y^*(0)}{f(0)^2} - v'\left(\frac{y^*(1)}{f(0)}\right)\frac{-y^*(1)}{f(0)^2}.$$

Using this approximation, we get that $dg/df(0) \leq 0$ if and only if

$$-v'\left(\frac{y^*(0)}{f(0)}\right)y^*(0) \leq -v'\left(\frac{y^*(1)}{f(0)}\right)y^*(1),$$

which is true for all $f(0)$, because $y^*(0) < y^*(1)$ for all $f(0)$ and $-v'$ is a positive and increasing function.

Strategy 4. Suppose that an agent decides to provide high effective labor $f(i^*)l^*(1)$ when his $\theta = 0$ and low effective labor $f(0)l^*(0)$ when his $\theta = 1$. Under this strategy, the agent's other choices at prices (r^*, w^*) and taxes (τ^*, ϕ^*), denoted by (c^4, k_1^4, i^4), are given as a solution of the following problem:

$$\max_{c_0,k_1,i,c_1(\theta)\geq 0} u(c_0) + \beta\left[\pi(0)\left(u(c_1(0)) + v\left(\frac{f(i^*)l^*(1)}{f(0)}\right)\right) + \pi(1)\left(u(c_1(1)) + v\left(\frac{f(0)l^*(0)}{f(i)}\right)\right)\right]$$

subject to

$$\begin{aligned} c_0 + i + k_1 &\leq k_0, \\ c_1(0) &\leq (1-\tau^*(1))r^*k_1 + w^*f(i^*)l^*(1) - \phi^*(1), \\ c_1(1) &\leq (1-\tau^*(0))r^*k_1 + w^*f(0)l^*(0) - \phi^*(0). \end{aligned}$$

The associated labor efforts are $l^4(0) = \frac{f(i^*)l^*(1)}{f(0)}$ and $l^4(1) = \frac{f(0)l^*(0)}{f(i^4)}$. By Inada conditions, $i^4 > 0$.

Recall from our analysis of strategy 2 that if an agent is to provide low effective labor $f(0)l^*(0)$ independently of the realization of his θ, then the individual choices (c^2, k_1^2, i^2, l^2) are a unique maximizer of his utility, and the level of utility attained is equal to the one of the allocation (c^*, k_1^*, i^*, l^*). Suppose now (ad absurdum) that the solution to the utility maximization problem under strategy

[10] This assumption is justified in a wide range of numerical examples we considered.

4, (c^4, k_1^4, i^4, l^4), yields strictly more utility than the allocation (c^*, k_1^*, i^*, l^*). This means that (c^4, k_1^4, i^4, l^4) yields strictly more utility than (c^2, k_1^2, i^2, l^2) does.

Consider now a plan of action in which the agent chooses c_0^4, k_1^4, i^4 at time 0 and then provides low effective labor independently of θ. This plan must yield strictly less utility than (c^2, k_1^2, i^2, l^2) because (c^2, k_1^2, i^2, l^2) is a unique maximizer under strategy 2. Under our ad absurdum assumption, (c^4, k_1^4, i^4, l^4) strictly dominates (c^2, k_1^2, i^2, l^2), which implies that choosing c_0^4, k_1^4, i^4 at time 0 and then providing low effective labor independently of θ is strictly dominated by (c^4, k_1^4, i^4, l^4). I.e.,

$$\begin{aligned} & u(c_0^4) + \beta\left(\pi(0)\left[u(c_1^4(0)) + v\left(\frac{f(i^*)l^*(1)}{f(0)}\right)\right] + \pi(1)\left[u(c_1^4(1)) + v\left(\frac{f(0)l^*(0)}{f(i^4)}\right)\right]\right) \\ > \quad & u(c_0^4) + \beta\left(\pi(0)\left[u(c_1^4(1)) + v\left(\frac{f(0)l^*(0)}{f(0)}\right)\right] + \pi(1)\left[u(c_1^4(1)) + v\left(\frac{f(0)l^*(0)}{f(i^4)}\right)\right]\right). \end{aligned}$$

This implies that

$$u(c_1^4(0)) + v\left(\frac{f(i^*)l^*(1)}{f(0)}\right) > u(c_1^4(1)) + v\left(\frac{f(0)l^*(0)}{f(0)}\right). \tag{71}$$

Since $f(0)l^*(0) < f(i^*)l^*(1)$, $f(i^4) > f(0)$ and $v', v'' < 0$, it follows from (71) that

$$u(c_1^4(0)) + v\left(\frac{f(i^*)l^*(1)}{f(i^4)}\right) > u(c_1^4(1)) + v\left(\frac{f(0)l^*(0)}{f(i^4)}\right). \tag{72}$$

The inequality (72) implies that

$$\begin{aligned} & u(c_0^4) + \beta\left(\pi(0)\left[u(c_1^4(0)) + v\left(\frac{f(i^*)l^*(1)}{f(i^4)}\right)\right] + \pi(1)\left[u(c_1^4(0)) + v\left(\frac{f(i)l^*(1)}{f(i^4)}\right)\right]\right) \\ > \quad & u(c_0^4) + \beta\left(\pi(0)\left[u(c_1^4(0)) + v\left(\frac{f(i^*)l^*(1)}{f(i^4)}\right)\right] + \pi(1)\left[u(c_1^4(1)) + v\left(\frac{f(0)l^*(0)}{f(i^4)}\right)\right]\right) \\ = \quad & u(c_0^4) + \beta\sum_{\theta\in\Theta}\left[\pi(\theta)u(c_1^4(\theta)) + v(l^4(\theta))\right], \end{aligned}$$

i.e., that choosing c_0^4, k_1^4, i^4 at date 0 and then providing high effective labor independently of θ strictly dominates allocation (c^4, k_1^4, i^4, l^4), and thus, by our ad absurdum assumption, also the optimal allocation (c^*, k_1^*, i^*, l^*). This is a contradiction because we have shown in our analysis of strategy 3 that providing high effective labor independently of θ cannot yield more utility than (c^*, k_1^*, i^*, l^*) for any choice of c_0, k_1, i. □

Proof of Proposition 4

Using the (37)–(39) expressions for optimal taxes, we get

$$\begin{aligned} E[1-\tau^*] &= \sum_{\theta\in\Theta}\pi(\theta)(1-\tau^*(\theta)) \\ &= \frac{\pi(0)u'(c_0^* + i^* - \widetilde{i}^*)}{r^*\beta u'(c_1^*(0))} + \frac{u'(c_0^*) - \pi(0)u'(c_0^* + i^* - \widetilde{i}^*)}{r^*\beta u'(c_1^*(1))}. \end{aligned} \tag{73}$$

The FOC (47) and (48) characterizing the optimal allocation imply that

$$\beta u'(c_1^*(0)) = \frac{(1+\alpha)}{(1-\alpha\frac{\pi(1)}{\pi(0)})}\beta u'(c_1^*(1)). \tag{74}$$

Substituting (74) into (73) yields

$$\begin{aligned} E[1-\tau^*] &= \frac{\pi(0)u'(c_0^*+i^*-\widetilde{i}^*)\left(1-\alpha\frac{\pi(1)}{\pi(0)}\right)}{r^*\beta u'(c_1^*(1))(1+\alpha)} + \frac{u'(c_0^*)-\pi(0)u'(c_0^*+i^*-\widetilde{i}^*)}{r^*\beta u'(c_1^*(1))} \\ &= \frac{(1+\alpha)u'(c_0^*)-\alpha u'(c_0^*+i^*-\widetilde{i}^*)}{(1+\alpha)r^*\beta u'(c_1^*(1))}. \end{aligned} \tag{75}$$

Using (22) to eliminate α from the denominator of (75), we get

$$E[1-\tau^*] = \frac{(1+\alpha)u'(c_0^*)-\alpha u'(c_0^*+i^*-\widetilde{i}^*)}{\frac{r^*\beta}{\sum\limits_{\theta\in\Theta}\frac{\pi(\theta)}{u'(c_1^*(\theta))}}} = 1,$$

where the last inequality follows directly from our modified Rogerson equation (21).

By the zero expected tax result shown above, in order to establish that $\tau^*(0) > 0 > \tau^*(1)$, all we need to show is $1-\tau^*(0) < 1-\tau^*(1)$. The fact that $i^* > \widetilde{i}^*$ gives us the following double inequality:

$$u'(c_0^*+i^*-\widetilde{i}^*) < u'(c_0^*) < u'(c_0^*)\left(\frac{1}{\pi(1)} - \frac{\pi(0)}{\pi(1)}\frac{u'(c_0^*+i^*-\widetilde{i}^*)}{u'(c_0^*)}\right),$$

which, together with $u'(c_1^*(1)) < u'(c_1^*(0))$ implies that

$$1-\tau^*(0) = \frac{u'(c_0^*+i^*-\widetilde{i}^*)}{r^*\beta u'(c_1^*(0))} < \frac{u'(c_0^*)\left(\frac{1}{\pi(1)} - \frac{\pi(0)}{\pi(1)}\frac{u'(c_0^*+i^*-\widetilde{i}^*)}{u'(c_0^*)}\right)}{r^*\beta u'(c_1^*(1))} = 1-\tau^*(1),$$

which completes the proof. $\square$

Proof of Proposition 5

Given that the expected marginal tax rate is zero both in the exogenous and in the endogenous skill model, it is enough to show that $1-\tau^*(0) < 1-\hat{\tau}(0)$. First, we show that $r^* > \hat{r}$. Using (in this order) the modified Rogerson condition (21), the facts that $i^* > \widetilde{i}^*$ and $\alpha > 0$, the assumption $\hat{c} = c^*$, and the standard Rogerson condition (42), we get

$$\begin{aligned} r^* &= E\left[\frac{u'(c_0^*)+\alpha\left(u'(c_0^*)-u'(c_0^*+i^*-\widetilde{i}^*)\right)}{\beta u'(c_1^*)}\right] \\ &> E\left[\frac{u'(c_0^*)}{\beta u'(c_1^*)}\right] = E\left[\frac{u'(\hat{c}_0)}{\beta u'(\hat{c}_1)}\right] = \hat{r}. \end{aligned}$$

Now, using $i^* > \widetilde{i^*}$, $r^* > \hat{r}$ and $\hat{c} = c^*$, we have

$$1 - \tau^*(0) = \frac{u'(c_0^* + i^* - \widetilde{i^*})}{r^* \beta u'(c_1^*(0))} < \frac{u'(c_0^*)}{r^* \beta u'(c_1^*(0))} < \frac{u'(\hat{c}_0)}{\hat{r} \beta u'(\hat{c}_1(0))} = 1 - \hat{\tau}(0),$$

which completes the proof. $\square$

References

Albanesi, S., (2005), Optimal Taxation of Entrepreneurial Capital Under Private Information, manuscript, Columbia University.

Albanesi, S., and Ch. Sleet, (2005), Dynamic optimal taxation with private information, forthcoming in the Review of Economic Studies.

Boldrin, M., and A. Montes, (2005), The Intergenerational State Education and Pensions, forthcoming in the Review of Economic Studies.

Davies, J., J. Zeng and J. Zhang, (2000), Consumption vs. income taxes when private human capital investments are imperfectly observable, Journal of Public Economics, Volume 77, Issue 1, 1-28.

United States Department of Treasury, Office of Tax Analysis, (2005), Treasury Fact Sheet on Savings and Investment, available at `http://www.taxreformpanel.gov/meetings/meeting-03162005.shtml`

Farhi, E. and I. Werning, (2005), Inequality, Social Discounting and Progressive Estate Taxation, NBER Working Paper W11408.

Golosov, M., N. Kocherlakota, and A. Tsyvinski, (2003), Optimal indirect and capital taxation, Review of Economic Studies, 70, 569-588.

Heckman, J. (1976), A Life-Cycle Model of Earnings, Learning, and Consumption, Journal of Political Economy, August 1976.

Jones, L.E., R.E. Manuelli, and P.E. Rossi, (1997), On the Optimal Taxation of Capital Income, Journal of Economic Theory, 73, 93-117.

Kapicka, M., (2005), Optimal Income Taxation and Human Capital Accumulation, manuscript UC Santa Barbara.

Kocherlakota, N., (2004), Wedges and taxes, American Economic Review, 94, 109-113.

Kocherlakota, N., (2005), Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation, Econometrica, 73, 1587–1621.

Kocherlakota, N., (2005a), Advances in dynamic optimal taxation, Minneapolis Fed working paper.

Mirrlees, J.A., (1971), An Exploration in the theory of optimum income taxation, Review of Economic Studies, 38, 175-208.

Palacios-Huerta, I., 2003, An empirical analysis of the risk properties of human capital returns, American Economic Review, 93, 948-964.

Palacios-Huerta, I., 2003a, Risk and Market Frictions as Determinants of the Human Capital Premium, Working paper, Brown University.

Rogerson, W.P., (1985), Repeated moral hazard, Econometrica, 53, 69-76.

Shaffer, H.G. (1961), Investment in Human Capital, American Economic Review, 51, 1026–1034.

Schultz, Th.W., (1961), Investment in Human Capital: Reply, American Economic Review, 51, 1-17.

Schultz, Th.W., (1961a), Investment in Human Capital: Comment, American Economic Review, 51, 1035–1039.