Optimal Taxation in a Life-Cycle Economy with Endogenous Human Capital Formation

Marek Kapička, UC Santa Barbara

Julian Neira, UC Santa Barbara

February 15, 2012

Abstract

We study efficient allocations and optimal policies in a life-cycle economy with risky human capital accumulation. The agents are ex-ante heterogeneous in their initial human capital and in their ability level. Ex-post, they also differ in their realization of shocks to human capital. The model incorporates two frictions. First, it assumes that ability and labor supply are both private information of the agents. Second, it adds a moral hazard component by assuming that schooling and realized rates of return to human capital are both private information. Since models with those three sources of heterogeneity are successful in replicating and explaining the distribution of earnings and consumption observed in the data, the model is thus both realistic enough to be useful for policy analysis, and tractable enough to carry out the analysis.

We assume that abilities are permanent and show that, under certain conditions, the inverse of the intratemporal wedge follows a random walk. This result is, to our knowledge, novel and implies that average intratemporal wedge increases over time.

We provide preliminary quantitative simulations for a two period economy and find that high ability agent face the largest expected increase in the intratemporal wedge.

J.E.L Codes: E6, H2

Keywords: optimal taxation, income taxation, human capital

1 Introduction

In this paper we explore the optimal tax structure and efficient allocations in a model where agents are heterogeneous in their ability to produce output, can invest in human capital to augment their productivity, and the rates of return to human capital evolve stochastically over one's lifetime. Abilities, labor supply and schooling decisions are all unobservable by the government. We (plan to) calibrate the economy to US data and solve numerically for the dynamics of efficient allocations and optimal tax schedules.

The main contribution of this paper is that it studies optimal taxation in a framework that is able to account for key features of the dynamics of the earnings and consumption that are observed in the data. As shown by Huggett, Ventura, and Yaron (2011), a properly parameterized life-cycle incomplete markets economy with risky human capital and heterogeneity in abilities is able to quantitatively account for the hump shaped profile of average earnings and an increase in the earnings dispersion and skewness over the life-cycle. Moreover, the stochastic process for earnings generated by the model is consistent with both leading statistical models, the RIP (restricted income profile) models (see e.g. MaCurdy (1982), Storesletten, Telmer, and Yaron (2004)) and the HIP (heterogeneous income profile) models (see e.g. Lillard and Weiss (1979), Guvenen (2007)).[1] Finally, the framework is also consistent with the increased dispersion in consumption over the life-cycle, as documented by Aguiar and Hurst (2008) or Primiceri and van Rens (2009). This research project takes the economy with risky human capital and heterogeneity in abilities as a starting point for the optimal taxation analysis.

The optimal tax problem combines a standard Mirrleesean private information friction arising from unobservability of individual abilities with a moral hazard friction arising from unobservability of schooling effort. The model has several notable properties. First, due to the moral hazard friction, consumption dispersion increases over the lifetime and the *inverse Euler equation holds*. Second, when the utility is additively separable in leisure and schooling we show that *the inverse of the intratemporal wedge follows a random walk*. This result is, to our knowledge, novel. It arises from two unique features of the model: the fact that the inverse Euler equation holds, and the fact that the agents receive no new information about

[1]The difference between RIP and HIP models is that in HIP models people face heterogeneous life-cycle earning profiles, while in RIP models individuals face similar life-cycle earning profiles.

abilities over the course of their lifetime. While the assumption of additive separability is, perhaps, not very realistic, the result is important as it serves as a useful benchmark for the analysis of the dynamics of intratemporal wedges.

A key aspect that makes the optimal tax problem tractable is that we extend the method developed in Boháček and Kapička (2008) (for riskless observable human capital) and Kapička (2008) (for riskless unobservable human capital). In both papers we show that with a first-order approach one can partially separate the redistributional dimension of the optimal tax problem, where the social planner redistributes resources across agents, and the dynamic dimension of the optimal tax problem, where the social planner chooses the optimal sequences of labor supply and schooling. In addition, the dynamic dimension can be conveniently written recursively. Using the recursive Lagrangean techniques of Marcet and Marimon (2009) we show that a similar decomposition is possible in this model. The result relies on the assumption that abilities are permanent (which is consistent with the model structure of Huggett, Ventura, and Yaron (2011)). The assumption that human capital is observable is also important for preserving tractability. It is, however, worth noting that due to unobservability of schooling the model shares some features typically associated with models with unobserved human capital, namely that the incentives to accumulate human capital must be provided indirectly, through the income taxes.

1.1 Relationship to the existing literature

Recent research on optimal taxation with private information followed the seminal contributions of Mirrlees (1971), Mirrlees (1976), and Mirrlees (1986), and extended them to dynamic economies. It has mostly focused on cases when the individual skills are exogenous (Golosov, Kocherlakota, and Tsyvinski (2003), Kocherlakota (2005)), Albanesi and Sleet (2006), Battaglini and Coate (2008), Farhi and Werning (2005), Werning (2007)). A most complete life-cycle analysis is Golosov, Tsyvinski, and Troshkin (2010) and Farhi and Werning (2010) who analyze optimal taxation in an environment with where individual skills are Markov (essentially a stripped-down version of the RIP model).

In contrast, this paper focuses on a case when individual skills are endogenous. A significant progress in this direction has been made by Grochulski and Piskorski (2005) who study a problem with unobservable risky human capital. However, investment in human capital is their model is only possible in the initial period and the dynamics in the remaining periods is technically similar to the above models with exogenous skills. Boháček and Kapička (2008) and Kapička (2008) study environments with riskless human capital in an infinite horizon setting. Boháček and Kapička (2008) assumes that human capital is observable, while Kapička (2008) assumes that it is unobservable.[2] While each of those models captures some important component of endogenous skill formation, neither of them is rich enough to fully capture the earnings and consumption dynamics observed in the data.

A related research has studied environments with risky physical capital: Albanesi (2007) studies a problem with observable risky physical capital, while Shourideh (2010) studies optimal allocation with entrepreneurial risk in a multiperiod setting.

2 The Model

Consider the following life-cycle economy. Agents live for $J > 1$ periods. They like to consume, dislike working and schooling, and have preferences given by

$$\mathbb{E}\sum_{j=1}^{J}\beta^{j-1}\left[U(c_j)-V(l_j,s_j)\right], \quad 0<\beta<1, \tag{1}$$

where j is age, c_j is consumption, l_j is labor, and s_j is schooling. The function U is strictly increasing, strictly concave, and differentiable. The function V is strictly increasing, strictly convex, and differentiable in both arguments.

The agent's earnings are determined by agent's ability a, current human capital h_j,

[2] Kapička (2006) analyzes the optimal steady state allocations in a similar environment with unobservable human capital and a restriction that the government can only use current income taxes and agents cannot borrow or save. See also Diamond and Mirrlees (2002) who analyze unobservable human capital investments in a static framework.

current labor l_j and the rental rate of human capital w:

$$y_j = wah_jl_j \tag{2}$$

The ability is constant over agent's lifetime and is known to the agents at the beginning of period 1. Ability and initial human capital h_1 are allowed to be correlated, and their joint distribution has density $q(a, h_1)$. The ability has a continuous support $A = (\underline{a}, \overline{a})$, with $\overline{a}$ possibly being infinite. Human capital in the first period, as well as in all other periods, has a continuous support $H = (\underline{h}, \overline{h})$, with $\overline{h}$ possibly infinite.

Human capital next period h_{j+1} depends on idiosyncratic human capital depreciation shock z_j, current human capital h_j, and on current schooling s_j:

$$h_{j+1} = e^{z_{j+1}}F(h_j, s_j) \tag{3}$$

where the function F is strictly increasing, strictly concave, and differentiable in both arguments. The idiosyncratic human capital shock is serially uncorrelated, but its density can depend on age j. As is standard in the moral hazard literature, it is useful to transform the state-space representation of the problem to work directly with the distribution induced over h_j. To that end, I construct a density function of human capital in period $j + 1$ conditional on $f_j = F(h_j, s_j)$, and denote it by $p_{j+1}(h_{j+1}|f_j)$. I also construct a density function of a sequence of human capital shocks $h^j = (h_1, \ldots, h_j)$ for a given history of schooling choices s^{j-1} and initial human capital h_1. It is given by

$$P^j(h^j|h_1, s^{j-1}) = p_2(h_2|F(h_1, s_1)) \ldots p_j(h_j|F(h_{j-1}, s_{j-1})) \quad j = 1, \ldots, J.$$

Finally, note that

$$\frac{\partial p_{j+1}\left[h_{j+1}, F(h_j, s_j)\right]}{\partial s_j} = \frac{\partial p_{j+1}\left[h_{j+1}, F(h_j, s_j)\right]}{\partial f}F_s(h_j, s_j) \tag{4}$$

This economy is identical to Huggett, Ventura, and Yaron (2011), with two exceptions. First, this model includes leisure. That is essential for thinking about optimal taxation. Second, the ability a affects earnings directly, rather than indirectly through the human capital production function. That is irrelevant in an incomplete markets economy studied by Huggett, Ventura, and Yaron (2011) if the human capital production function takes the Ben-Porath form:

$$F(h,s) = h + (hs)^{\alpha}. \tag{5}$$

To see that both formulations are isomorphic, redefine human capital as follows: Let $\tilde{h} = ha$ and $\tilde{a} = a^{1-\alpha}$. Then the law of motion for human capital is $\tilde{F}(h,s) = \tilde{h} + \tilde{a}(\tilde{h}s)^{\alpha}$, and the earnings are $y = w\tilde{h}l$, identical to the ones in Huggett, Ventura, and Yaron (2011). However, both formulations have different implications in a Mirrleesean economy with private information and observable human capital where it makes a difference whether h or ha is observed. The formulation chosen in this research proposal has the advantage that it is entirely consistent with the existing optimal taxation literature.[3]

3 An Incomplete Markets Economy

Consider first an economy where the agents can self-insure against idiosyncratic shocks by saving. Denote the saving made in period j by k_{j+1}. In addition, the government taxes the agents of age j by an exogenously given tax function $T_j(e_j, k_j)$. The savings yield a gross rate of return r. The economy is essentially a Bewley-type incomplete markets economy.

In a market economy, the agent with initial human capital h_1 and ability a choose sequences of labor supply, schooling and human capital $\{c_j(a,h_1,z^j), l_j(a,h_1,z^j), s_j(a,h_1,z^j),$

[3] Both formulations are again isomorphic if both h and a are either observable or unobservable. The first case is inconsistent with the Mirrleesean framework, while the second one would be extremely hard to solve in general.

$h_{j+1}(a, h_1, z^{j+1})\}$ to maximize the lifetime utility

$$\sum_{j=1}^{J} \beta^{j-1} \int_{Z^j} \left[U\left(c_j(a, h_1, z^j)\right) - V\left(l_j(a, h_1, z^j), s_j(a, h_1, z^j)\right)\right] \hat{P}^j(h_1, z^j)\, dz^j, \tag{6}$$

subject to the budget constraint

$$c_j(a, h_1, z^j) + k_{j+1}(a, h_1, z^j) = rk_j(a, h_1, z^j) + y_j(a, h_1, z^j) - T_j\left[y_j(a, h_1, z^j), k_j(a, h_1, z^j)\right],$$

law of motion for human capital (3), the earnings function (2), and a terminal condition $k_{J+1} = 0$. They take their ability and their initial human capital as given.

A recursive formulation is as follows. The value function $G_j(h, k; a)$ satisfies

$$G_j(h, k; a) = \max_{c,l,s,k'} \left\{U(c) - V(l, s) + \beta \int_{H_{j+1}} G_{j+1}(h', k'; a) p_{j+1}(h'|F(h, s))\, dh'\right\}, \quad j = 1, \ldots, J$$

subject to

$$c + k' = rk + awhl - T_j\,(awhl, k)$$

with the terminal condition $G_{J+1} = 0$.

4 Optimal Taxation in a J Period Model

4.1 Efficient Allocations

The information structure is as follows: ability a, labor supply l_j and schooling s_j are all private information. Consumption c_j, earnings y_j and human capital realizations h_j are all publicly observed. Let $s_j(h^j)$ be schooling in period j after a history of human capital realizations h^j and let $s = \{s_j(h^j)\}_{j=1}^{J}$ be an arbitrary state contingent schooling plan. Define the utility of type a agent that reports $\hat{a}$ and chooses schooling plan s by

$$\hat{W}(\hat{a}, s|a, h_1) = \sum_{j=1}^{J} \beta^{j-1} \int_{H^{j-1}} \left[U(c_j(\hat{a}, h^j)) - V\left(\frac{y_j(\hat{a}, h^j)}{ah_j}, s_j(h^j)\right)\right] P^j\left(h^j|h_1, s^{j-1}(h^{j-1})\right) dh^j,$$

where consumption plan $c(a) = \{c_j(a, h^j)\}_{j=1}^J$ and earnings plan $y(a) = \{y_j(a, h^j)\}_{j=1}^J$ are allocations chosen by the social planner. Let $\hat{s}(\hat{a}|a, h_1)$ be the utility maximizing schooling plan that an $a-$type would choose if he reported $\hat{a}$. Define also the utility maximizing schooling plan conditional on truthtelling by $s(a, h_1) = \hat{s}(a|a, h_1)$, and let $W(a) = \hat{W}(a, s(a, h_1)|a, h_1)$ be the corresponding lifetime utility. Incentive compatibility requires that the agent prefers to tell the truth about her ability and that the schooling choice maximizes his utility:

$$W(a, h_1) \geq \hat{W}(\hat{a}, \hat{s}(\hat{a}|a, h_1)|a, h_1) \quad \forall a, \hat{a} \in A \forall h_1 \in H. \tag{7}$$

The planner chooses consumption plan $c(a)$, earnings plan $y(a)$ and schooling plan $\hat{s}(\hat{a}|a, h_1)$ to maximize the expected lifetime utility[4]

$$\int_{A\times H} W(a, h_1) q(a, h_1)\, da\, dh_1 \tag{8}$$

subject to the incentive constraint (7) and the resource constraint

$$\int_{A\times H} \sum_{j=1}^{J} r^{-j+1} \int_{H^{j-1}} \left[c_j(a, h^j) - y_j(a, h^j)\right] P^j\left(h^j|h_1, s^{j-1}(a, h^{j-1})\right) dh^j\, q(a, h_1)\, da\, dh_1 \leq 0. \tag{9}$$

To reduce the complexity of the problem, we will assume that $r = \beta^{-1}$.

[4]We plan to solve the problem for other social welfare functions, namely the Rawlsian social welfare function as well.

4.1.1 First-Order Approach

The first-order approach replaces the incentive constraint (7) with two conditions. The first one is the first-order condition in schooling and says that, at the optimum, the marginal costs of schooling (given by the disutility from spending an additional unit of time by schooling) must be equal to the expected marginal benefit of schooling (given by the additional utility arising from the fact that the distribution of future human capital shocks is now more favorable):

$$V_s\left(\frac{y_j(a,h^j)}{ah_j}, s_j(a,h^j)\right) = \sum_{i=1}^{J-j} \beta^i \int_{H^i} \left[U(c_{j+i}(a,h^j,\zeta^i)) - V\left(\frac{y_{j+i}(a,h^j,\zeta^i)}{a\zeta_i}, s_{j+i}(a,h^j,\zeta^i)\right)\right] \times \frac{P_{s_j}{}^{j+i}\left(h^j,\zeta^i|h_1, s^{j+i-1}(a,h^j,\zeta^{i-1})\right)}{P^j\left(h^j|h_1, s^{j-1}(a,h^{j-1})\right)} d\zeta^i \tag{10}$$

for all histories $h^j \in H^j$. The second one is an envelope condition saying how the lifetime utility needs to vary with ability in order to deter the agent from misreporting his type:

$$W(a,h_1) = \underline{W}(h_1) + \int_{\underline{a}}^{a} \sum_{j=1}^{J} \beta^{j-1} \int_{H^{j-1}} \left[V_l\left(\frac{y_j(\varepsilon,h^j)}{\varepsilon h_j}, s_j(\varepsilon,h^j)\right) \frac{y_j(\varepsilon,h^j)}{\varepsilon h_j} P^j(h^j|h_1, s^{j-1}(\varepsilon,h^{j-1}))\, dh^j\right] \frac{d\varepsilon}{\varepsilon} \tag{11}$$

where $\underline{W}(h_1) = W(\underline{a}, h_1)$ is the lifetime utility of the least able agent with initial human capital h_1. For now, I will assume that the first-order approach is valid and the set of allocations that satisfy (7) is identical to the set of allocations that satisfy (10) and (11). We will return to the problem of validity of the first-order approach later.

Replacing the incentive constraint (7) with (10) and (11) leads to a *relaxed* planning problem. The relaxed planning problem maximizes (8) subject to the constraints (9), (11) and (10) by choosing c, y, s and $\underline{W}$. Let λ, $\theta(a,h_1)q(a,h_1)$ and $\beta^{j-1}\phi_j(a,h^j)q(a,h_1)$, $j = 1,\ldots,J-1$ be the (appropriately normalized) Lagrange multipliers on the resource constraint (9), on the envelope condition (11), and on the first order condition (10). The Lagrangean is

$$
\mathcal{L} = \int_{A\times H} \sum_{j=1}^{J} \beta^{j-1} \int_{H^{j-1}} \Bigg\{ U(c_j(a,h^j)) - V\left(\frac{y_j(a,h^j)}{ah_j}, s_j(a,h^j)\right) - \lambda \left[c_j(a,h^j) - y_j(a,h^j)\right]
$$

$$
+ \theta(a,h_1)\left[U(c_j(a,h^j)) - V\left(\frac{y_j(a,h^j)}{ah_j}, s_j(a,h^j)\right) - \int_{\underline{a}}^{a} V_l\left(\frac{y_j(\varepsilon,h^j)}{\varepsilon h_j}, s_j(\varepsilon,h^j)\right)\frac{y_j(\varepsilon,h^j)}{\varepsilon h_j}\frac{d\varepsilon}{\varepsilon}\right]
$$

$$
- \phi_j(a,h^j)\left[\frac{V_s\left(\frac{y_j(a,h^j)}{ah_j}, s_j(a,h^j)\right)}{F_s\left(h_j, s_j(a,h^j)\right)} - \sum_{i=1}^{J-j}\beta^i \int_{H^i}\left[U(c_{j+i}(a,h^j,\zeta^i)) - V\left(\frac{y_{j+i}(a,h^j,\zeta^i)}{a\zeta_i}, s_{j+i}(a,h^j,\zeta^i)\right)\right]\right.
$$

$$
\left.\times \frac{P_{s_j}{}^{j+i}(h^j,\zeta^i|h_1, s^{j+i-1}(a,h^j,\zeta^{i-1}))}{P^j(h^j|h_1, s^{j-1}(a,h^{j-1}))}\frac{1}{F_s\left(h_j, s_j(a,h^j)\right)} d\zeta^i\right]
$$

$$
- \theta(a,h_1)\underline{W}(h_1)\Bigg\} P^j(h^j|h_1, s^{j-1}(a,h^{j-1}))\, dh^j\, q(a,h_1)\, da\, dh_1
$$

The first-order condition in $\underline{W}(h_1)$ implies

$$
\int_A \theta(a,h_1) q(a,h_1)\, da = 0 \quad \forall h_1 \in H. \tag{12}
$$

Integrating the Lagrangean by parts and using (12), one obtains

$$
\mathcal{L} = \int_{A\times H} \sum_{j=1}^{J} \beta^{j-1} \int_{H^{j-1}} \Bigg\{ (1+\theta(a,h_1))\left[U(c_j(a,h^j)) - V\left(\frac{y_j(a,h^j)}{ah_j}, s_j(a,h^j)\right)\right]
$$

$$
- \lambda\left[c_j(a,h^j) - y_j(a,h^j)\right]
$$

$$
- \phi_j(a,h^j)\left[\frac{V_s\left(\frac{y_j(a,h^j)}{ah_j}, s_j(a,h^j)\right)}{F_s\left(h_j, s_j(a,h^j)\right)} - \sum_{i=1}^{J-j}\beta^i \int_{H^i}\left[U(c_{j+i}(a,h^j,\zeta^i)) - V\left(\frac{y_{j+i}(a,h^j,\zeta^i)}{a\zeta_i}, s_{j+i}(a,h^j,\zeta^i)\right)\right]\right.
$$

$$
\left.\times \frac{P_{s_j}{}^{j+i}(h^j,\zeta^i|h_1, s^{j+i-1}(a,h^j,\zeta^{i-1}))}{P^j(h^j|h_1, s^{j-1}(a,h^{j-1}))}\frac{1}{F_s\left(h_j, s_j(a,h^j)\right)} d\zeta^i\right]
$$

$$
- \Theta(a,h_1) V_l\left(\frac{y_j(a,h^j)}{ah_j}, s_j(a,h^j)\right)\frac{y_j(a,h^j)}{ah_j}\Bigg\} P^j(h^j|h_1, s^{j-1}(a,h^{j-1}))\, dh^j\, q(a,h_1)\, da\, dh_1
$$

where $\Theta(a,h_1)$ is the cross-sectional cumulative of the Lagrange multipliers on the envelope condition:

$$\Theta(a, h_1) = \frac{1}{aq(a)} \int_a^{\overline{a}} \theta(\varepsilon, h_1) q(\varepsilon, h_1)\, d\varepsilon. \tag{13}$$

Finally, rearranging the terms involving ϕ, one obtains

$$\begin{aligned}
\mathcal{L} = \int_{A\times H} \sum_{j=1}^{J} \beta^{j-1} \int_{H^{j-1}} \Bigg\{ & \left(1 + \theta(a, h_1) + \Phi_j(a, h^j)\right) \left[U(c_j(a, h^j)) - V\left(\frac{y_j(a, h^j)}{ah_j}, s_j(a, h^j)\right)\right] \\
& - \lambda\left[c_j(a, h^j) - y_j(a, h^j)\right] - \phi_j(a, h^j) \frac{V_s\left(\frac{y_j(a,h^j)}{ah_j}, s_j(a, h^j)\right)}{F_s\left(h_j, s_j(a, h^j)\right)} \\
& - \Theta(a, h_1) V_l\left(\frac{y_j(a, h^j)}{ah_j}, s_j(a, h^j)\right) \frac{y_j(a, h^j)}{ah_j} \Bigg\} P^j(h^j | h_1, s^{j-1}(a, h^{j-1}))\, dh^j\, q(a, h_1)\, da\, dh_1
\end{aligned}$$

where $\Phi_j(a, h^j)$ is the intertemporal cumulative of the Lagrange multipliers on the first order condition in schooling: $\Phi_1(a, h_1) = 0$ and then for $j = 2, \ldots, J$,

$$\Phi_j(a, h^j) = \sum_{i=1}^{j-1} \phi_i(a, h^i) \frac{p_{f\,i+1}\left(h_{i+1} | F(h_i, s_i(a, h^i))\right)}{p_{i+1}\left(h_{i+1} | F(h_i, s_i(a, h^i))\right)}.$$

The expression for Φ is simplified by the property (4). Note that both cumulative multipliers have an economic interpretation: $\Theta(a, h_1)$ indicates how much the planner desires to redistribute resources across agents, and is a key in determining how much to distort labor supply of a $a-$type agent. $\Phi_j(a, h^j)$ indicates how costly it is for the social planner to respect the first-order condition in schooling (10). Note that $\Phi_1(a, h_1) = 0$.

The planning problem an now be written as a saddle point of the Lagrangean:

$$\max_{c,y,s} \min_{\lambda,\theta,\phi} \mathcal{L}.$$

4.2 Theoretical Implications

The first-order conditions in consumption are

$$\frac{1}{U'(c_j(a,h^j))} = \frac{1+\theta(a,h_1)+\Phi_j(a,h^j)}{\lambda} \tag{14}$$

This implies that, conditional on ability, there is a dispersion in consumption in all the periods except for the first one. In addition, the Inverse Euler Equation holds, as the next Proposition shows:

Proposition 1

$$\frac{1}{U'(c_j(a,h^j))} = \int_H \frac{1}{U'(c_{j+1}(a,h^{j+1}))} p_{j+1}\left(h_{j+1}|F(h_j,s_j(a,h^j))\right)\, dh_{j+1} \quad \forall h^j \in H^j$$

Proof. Note that

$$\Phi_{j+1}(a,h^{j+1}) = \Phi_j(a,h^j) + \phi_j(a,h^j)\frac{p_{f\,j+1}\left(h_{j+1}|F(h_j,s_j(a,h^j))\right)}{p_{j+1}\left(h_{j+1}|F(h_j,s_j(a,h^j))\right)}.$$

Hence

$$\int_H \Phi_{j+1}(a,h^{j+1}) p_{j+1}\left(h_{j+1}|F(h_j,s_j(a,h^j))\right)\, dh_{j+1} = \Phi_j(a,h^j).$$

The rest of the proof follows from the first-order condition (14). ■

The first-order conditions in labor imply that

$$\begin{aligned}\frac{ah_j}{V_l\left(l_j(a,h^j),s_j(a,h^j)\right)} - \frac{1}{U'(c_j(a,h^j))} &= \left[1+\gamma\left(l_j(a,h^j),s_j(a,h^j)\right)\right]\frac{\Theta(a)}{\lambda} \\ &+ \frac{\phi_j(a,h^j)}{\lambda}\frac{V_{ls}\left(l_j(a,h^j),s_j(a,h^j)\right)}{V_l\left(l_j(a,h^j),s_j(a,h^j)\right)}\frac{1}{F_s\left(h_j,s_j(a,h^j)\right)}.\end{aligned}$$

where $\gamma(l_j, s_j) = \frac{l_j V_{ll}(l_j,s_j)}{V_l(l_j,s_j)}$. Define the intratemporal wedges by

$$\tau_j(a, h^j) = 1 - \frac{V_l\left(l_j(a,h^j), s_j(a,h^j)\right)}{ah_j U'(c_j(a,h^j))}.$$

The first-order conditions in labor imply that the wedges satisfy

$$\frac{1}{U'(c_j(a,h^j))}\frac{\tau_j(a,h^j)}{1-\tau_j(a,h^j)} = \left[1+\gamma\left(l_j(a,h^j), s_j(a,h^j)\right)\right]\frac{\Theta(a)}{\lambda} + \frac{\phi_j(a,h^j)}{\lambda}\frac{V_{ls}\left(l_j(a,h^j), s_j(a,h^j)\right)}{V_l\left(l_j(a,h^j), s_j(a,h^j)\right)}\frac{1}{F_s\left(h_j, s_j(a,h^j)\right)}.$$

If labor supply and schooling are additively separable and labor supply has constant elasticity then we have the following sharp characterization of the intratemporal wedges:

Proposition 2 *Suppose that $V(l,s) = \frac{l^{1+\gamma}}{1+\gamma} + g(s)$. Then*

$$\frac{1}{\tau_j(a,h^j)} = \int_H \frac{1}{\tau_{j+1}(a,h^{j+1})} p_{j+1}\left(h_{j+1}|F(h_j, s_j(a,h^j))\right) dh_{j+1}.$$

Proof. The expressions for wedges imply that

$$\frac{1}{U'(c_j(a,h^j))}\frac{\tau_j(a,h^j)}{1-\tau_j(a,h^j)} = \frac{1}{U'(c_{j+1}(a,h^{j+1}))}\frac{\tau_{j+1}(a,h^{j+1})}{1-\tau_{j+1}(a,h^{j+1})}.$$

Rearranging,

$$\frac{\tau_j(a,h^j)}{1-\tau_j(a,h^j)}\frac{1-\tau_{j+1}(a,h^{j+1})}{\tau_{j+1}(a,h^{j+1})} = \frac{U'(c_j(a,h^j))}{U'(c_{j+1}(a,h^{j+1}))}.$$

Since (1) holds,

$$\frac{1-\tau_j(a,h^j)}{\tau_j(a,h^j)} = \int_H \frac{1-\tau_{j+1}(a,h^{j+1})}{\tau_{j+1}(a,h^{j+1})} p_{j+1}\left(h_{j+1}|F(h_j, s_j(a,h^j))\right) dh_{j+1}.$$

Rearranging, the result follows. ■

The result is due to several facts. First, the tax revenue of an $a-$type agent is proportional to $\frac{\tau_j(a,h^j)}{1-\tau_j(a,h^j)}$ (Saez (2001)). Second, if the assumptions of Proposition 2 hold then (since the ability shock is permanent) the social planner wants to keep the tax revenue valued at the utility cost $\frac{1}{U'(c_j(a,h^j))}$ constant over time and state. Hence the expression $\frac{1}{U'(c_j(a,h^j))}\frac{\tau_j(a,h^j)}{1-\tau_j(a,h^j)}$ is constant over time and state. Since $\frac{1}{U'(c_j(a,h^j))}$ follows a random walk, the result follows. Jensen's inequality then implies that the average intratemporal wedge is increasing over time,

$$\tau_j(a,h^j) < \int_H \tau_{j+1}(a,h^{j+1}) p_{j+1}\left(h_{j+1}|F(h_j,s_j(a,h^j))\right) dh_{j+1}.$$

and so the intratemporal wedge is on average increasing over time. Second (this needs more work), since the intratemporal wedge is between zero and one, in an infinite horizon economy the intratemporal wedge will converge to one.

4.3 A Recursive formulation

The above problem can be solved using the following three-step procedure adapted from Boháček and Kapička (2008) and Kapička (2008). First, fix a Lagrange multiplier λ. Second, fix a function $\theta(a,h_1)$ and compute the cumulative Lagrange multiplier $\Theta(a,h_1)$ using (13). Conditional on those values, the problem has a recursive representation that uses the recursive Lagrangean method of Marcet and Marimon (2009). Denote Φ to be the costate variable corresponding to the cumulative Lagrange multiplier on the first-order condition (10) in schooling. Let $\Omega_j(h,\Phi;a,h_1)$ be the value of having human capital h and a costate variable Φ at the beginning of period j for an agent with ability a and initial human capital h_1. It can be shown that Ω_j is given by

$$\begin{aligned}\Omega_j(h,\Phi;a,h_1) = \max_{c,l,s}\min_{\phi}\Bigg\{ &[1+\theta(a,h_1)+\Phi]\left[U(c)-V(l,s)\right] - \lambda(c-ahl) - \phi\frac{V_s(l,s)}{F_s(h,s)} \\ &- V_l(l,s)\, l\Theta(a,h_1) + \beta\int_H \Omega_{j+1}\left[h',\Phi'(h');a,h_1\right] p_{j+1}(h'|F(h,s))\,dh'\Bigg\}\end{aligned} \quad (15)$$

where the law of motion for Φ is

$$\Phi'(h') = \Phi + \phi\frac{p_{f\,j+1}(h'|F(h,s))}{p_{j+1}(h'|F(h,s))}.$$

The dynamic program is initiated at $(h, \Phi) = (0, h_1)$, and is terminated with $\Omega_{J+1} = 0$.

Once the problem is computed, one updates the θ function by using a period 1 first order condition in consumption. This generates a new function $T\theta$:

$$T\theta(a, h_1) = \frac{\lambda}{U'\left(c_1(a, h_1)\right)} - 1$$

The iterations proceed until T converges. Finally, one iterates on the Lagrange multiplier λ until the resource constraint (9) clears.

The dynamic program (15) is clearly in the heart of this computational procedure. It has four state variables, but it is worth noting that only two of them (h and Φ) are changing over time; the other two (a and h_1) are constant. Also, h_1 enters the problem only through θ and Θ, which will somewhat simplify the computational procedure. Furthermore, the optimization problem itself is relatively simple, as it features no constraints (apart from nonnegativity constraints on c, l and s) and only four variables. It is therefore expected that, despite its complexities, it will be feasible to solve the problem numerically.[5]

Note also that the dynamic program (15) makes it easy to pinpoint the contribution of both frictions that are present in the model. In the absence of the moral hazard friction one would set $\Phi = \phi = 0$. In the absence of the private information friction one would set $\theta = \Theta = 0$. The dynamic program can therefore be easily adapted to "shut down" either one of those frictions to study its contribution to the optimal tax problem.

[5]We have solved a two period version of the dynamic program (15) with relative ease.

5 A Two Period Example

We now solve for the Pareto efficient allocation in a simple two period example. The utility functions are given by

$$U(c) = \frac{c^{1-\rho}}{1-\rho}$$
$$V(l,s) = \frac{l^{1+\gamma}}{1+\gamma} + \frac{s^{1+\epsilon}}{1+\epsilon}.$$

The human capital production function takes the Ben-Porath form (5). I assume that the ability a is lognormally distributed with density $q(a) = LN(\mu_a, \sigma_a^2)$. Human capital shocks h_2 are drawn from a truncated lognormal distribution, $p(h_2|f) = LN(\mu_z + \ln f, \sigma_z^2)$. A simple differentiation yields that $p_f(h_2|f) = \frac{\ln h_2 - \mu_z - \ln f}{\sigma_z^2} \frac{p(h_2|f)}{f}$.

We make two simplifying assumptions in this numerical example. First, we assume that the initial human capital h_1 is the same for everyone. Second, we modify the general setup by assuming that the social planner faces a Rawlsian problem of maximizing the utility of the least able agent. This implies that $\theta(a) = 1$ and that $\Theta(a) = \frac{1-Q(a)}{aq(a)}$.

The first-order conditions in the second period determine the optimal consumption and labor supply as a function of a and λ:

$$\frac{1+\theta(a)}{\lambda} + \phi \frac{\ln h - \mu_z - \ln f}{\sigma_z^2} = c_2(a,h)^{\rho}$$
$$\frac{1+\theta(a)}{\lambda} + \phi \frac{\ln h - \mu_z - \ln f}{\sigma_z^2} = ahl_2(a,h)^{-\gamma} - \frac{1+\gamma}{aq(a)}\Theta(a)$$

For the numerical simulations we assume that $\rho = 1$, $\gamma = 2$ and $\epsilon = 2$. That is, the Frisch elasticity of labor supply is 0.5, and the elasticity of schooling is also 0.5. The discount rate is $\beta = 0.96$. The initial human capital is $h_1 = 1$ and the human capital production function has $\alpha = 1$. The parameters of the ability distribution are $\sigma_a = 0.5$ and $\mu_a = -\frac{\sigma_a^2}{2}$, implying that the ability distribution has mean one. The parameters of the human capital shock distribution are $\sigma_z = 1$ and $\mu_z = -\frac{\sigma_z^2}{2}$, implying again that the shocks have mean one.

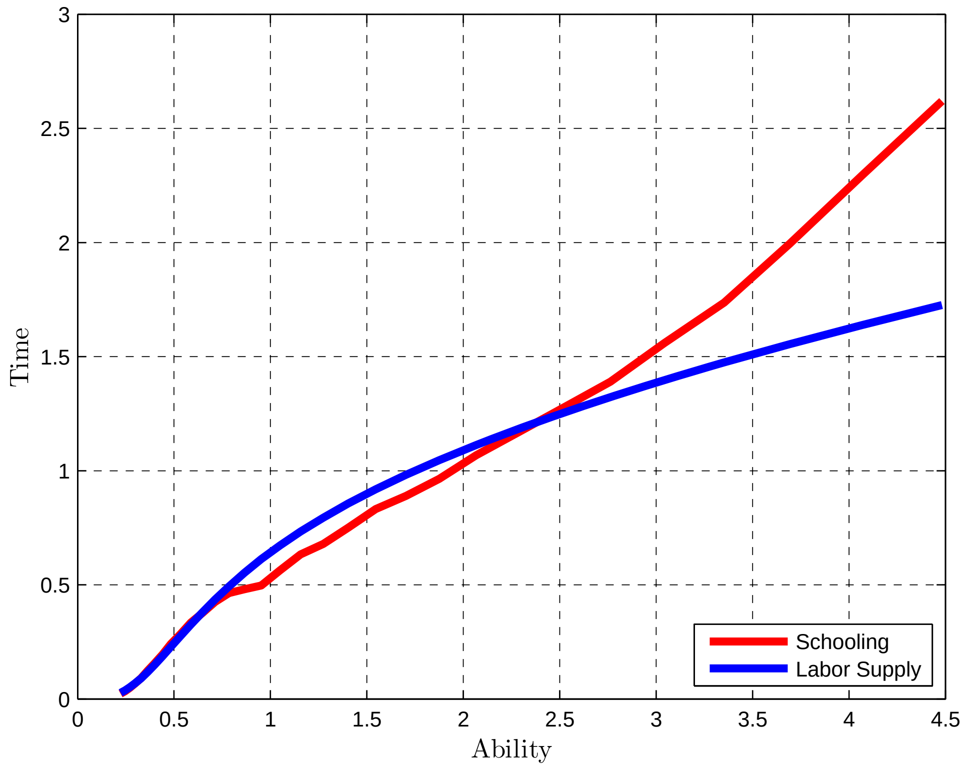

Figure 1: Optimal schooling in the first period

Since human capital shocks have mean one, human capital decreases on average over time for people with schooling less than one their, and increases on average for people with schooling more than one. Figure 1 shows that schooling increases with ability, but for majority of people human capital will be lower in the second period. Labor supply also increases with ability. (The fact that labor supply and schooling are not substitutes is in part a consequence of the particular functional form of V, and will not likely hold more generally, see Kapička (2008)).

The intratemporal wedge in the first period and the expected intratemporal wedge in the second period are shown in Figure 2. The figure confirms that the expected intratemporal wedge is higher than the intratemporal wedge in the first period. This confirms the theoretical findings in Proposition 2. The difference is most pronounced for higher ability level where the agents expect next period intratemporal wedge to be about 10% higher than

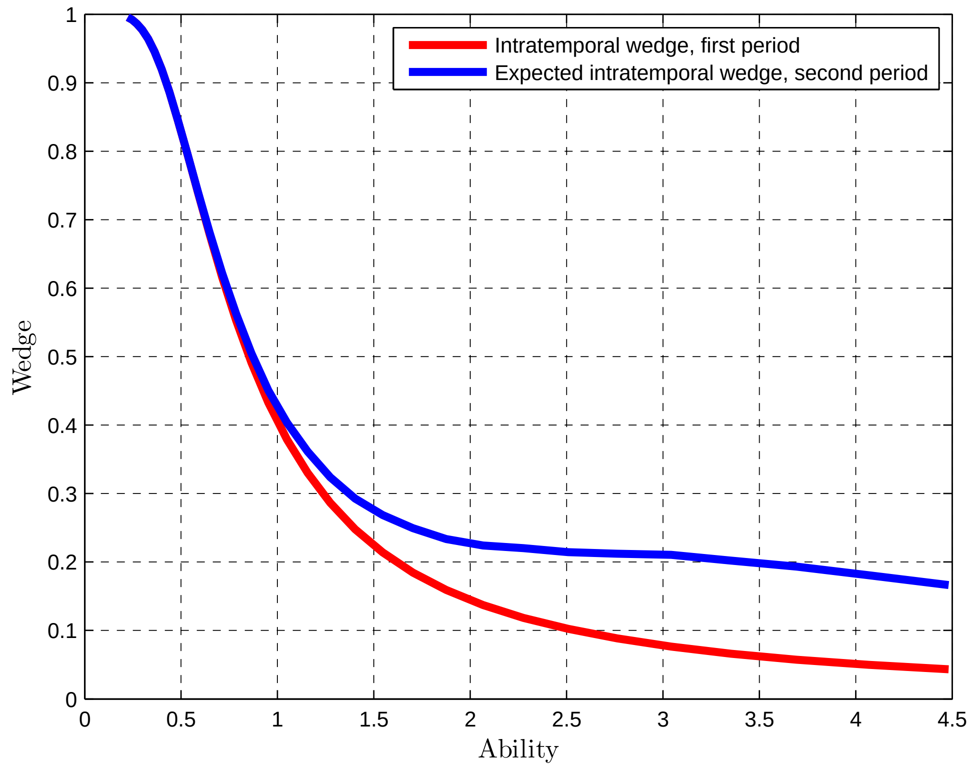

Figure 2: Intratemporal Wedges

today. Overall, the intratemporal wedge in period 1 decreases with abilities and converges to zero due to the fact that abilities are lognormally distributed.

Figure 3 shows the intratemporal wedge in the second period as a function of human capital. The wedges are shown for three selected ability levels, low, medium and high. In all cases the intratemporal wedge is very high for low human capital realizations and then decreases with human capital. The decrease is most rapid for higher ability levels.

6 Calibration of the J Period Model

To be completed.

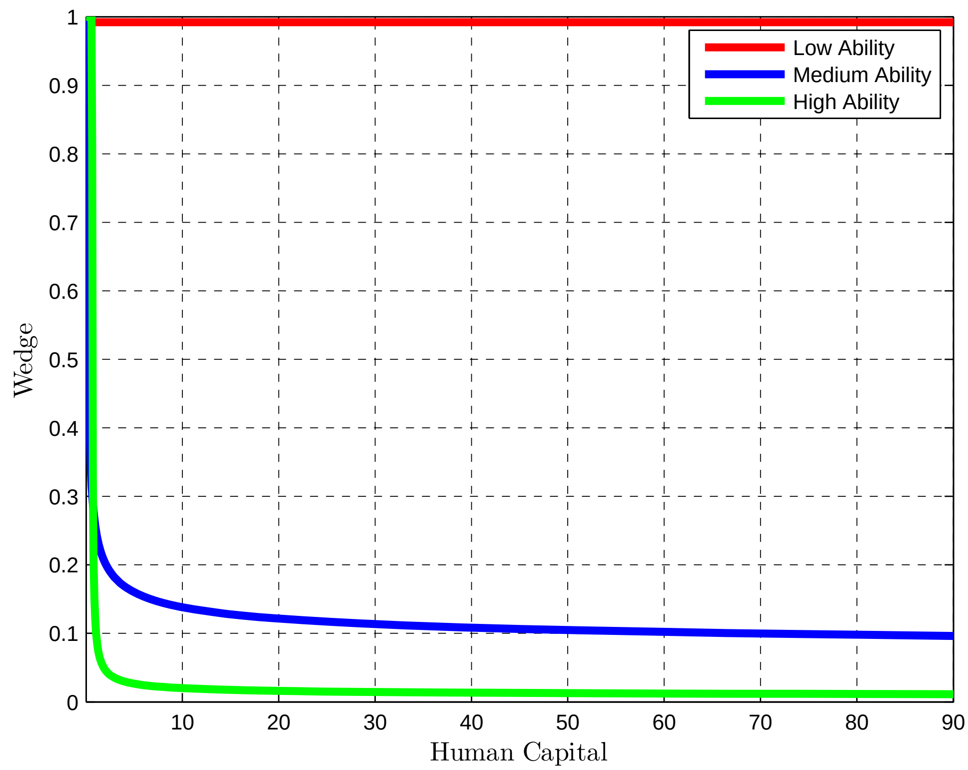

Figure 3: Intratemporal Wedge in the Second Period

7 Quantitative Results for the J Period Model

To be completed.

8 Conclusions

To be completed.

References

Aguiar, M. and E. Hurst (2008). Deconstructing lifecycle expenditure. Working Paper 13893, NBER. 2

Albanesi, S. (2007). Optimal taxation of entrepreneurial capital with private information. Working paper, Columbia University. 4

Albanesi, S. and C. Sleet (2006). Dynamic optimal taxation with private information. *The Review of Economic Studies 73*(1), 1–30. 3

Battaglini, M. and S. Coate (2008). Pareto efficient income taxation with stochastic abilities. *Journal of Public Economics 92*(3-4), 844–868. 3

Boháček, R. and M. Kapička (2008). Optimal human capital policies. *Journal of Monetary Economics 55*, 1–16. 3, 4, 14

Diamond, P. and J. Mirrlees (2002). Optimal taxation and the le chatelier principle. Working paper, MIT. 4

Farhi, E. and I. Werning (2005). Inequality and social discounting. *Journal of Political Economy 115*(1), 365–402. 3

Farhi, E. and I. Werning (2010). Insurance and taxation over the life-cycle. Working paper, MIT. 4

Golosov, M., N. R. Kocherlakota, and A. Tsyvinski (2003). Optimal indirect and capital taxation. *The Review of Economic Studies 70*, 569–587. 3

Golosov, M., A. Tsyvinski, and M. Troshkin (2010). Optimal dynamic taxes. Working paper, Yale University. 4

Grochulski, B. and T. Piskorski (2005). Optimal wealth taxes with risky human capital. Working paper, Federal Reserve Bank of Richmond. 4

Guvenen, F. (2007). Learning your earning: Are labor income shocks really very persistent? *American Economic Review 97*(3), 687–712. 2

Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *American Economic Review 101*(7), 2923–54. 2, 3, 6

Kapička, M. (2006). Optimal taxation and human capital accumulation. *Review of Economic Dynamics 9*, 612–639. 4

Kapička, M. (2008). The dynamics of optimal taxation when human capital is endogenous. Working paper, UC Santa Barbara. 3, 4, 14, 17

Kocherlakota, N. R. (2005). Zero expected wealth taxes: A Mirrless approach to dynamic optimal taxation. *Econometrica 73*, 1587–1621. 3

Lillard, L. A. and Y. Weiss (1979). Components of variation in panel earnings data: American scientists 1960-70. *Econometrica 47*, 437–453. 2

MaCurdy, T. (1982). The use of time-series processes to model the er- ror structure of earnings in a longitudinal data analysis. *Journal of Econometrics, 18*, 83–114. 2

Marcet, A. and R. Marimon (2009). Recursive contracts. *Unpublished manuscript*. 3, 14

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies 38*, 175–208. 3

Mirrlees, J. A. (1976). Optimal tax theory: A synthesis. *Journal of Public Economics 6*, 327–358. 3

Mirrlees, J. A. (1986). The theory of optimal taxation. In K. J. Arrow and M. D. Intriligator (Eds.), *Handbook of Mathematical Economics, vol. III, Chapter 24*. Elsevier. 3

Primiceri, G. E. and T. van Rens (2009). Heterogeneous life-cycle profiles, income risk and consumption inequality. *Journal of Monetary Economics 56*, 20–39. 2

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The Review of Economic Studies 68*, 205–229. 13

Shourideh, A. (2010). Optimal taxation of entrepreneurial income: A mirrleesian approach to capital accumulation. Working paper, University of Minnesota. 4

Storesletten, K., C. Telmer, and A. Yaron (2004). Consumption and risk sharing over the life cycle. *Journal of Monetary Economics 51*, 609–633. 2

Werning, I. (2007). Optimal fiscal policy with redistribution. *Quarterly Journal of Economics 22*(3), 925967. 3