Human Capital and Optimal Redistribution

Winfried Koeniger* and Julien Prat**

April 2014

Preliminary and incomplete
(Please do not circulate)

Abstract

We show that more human capital improves incentives in a standard optimal taxation problem: common assumptions about preferences and technology imply that the disutility of labor decreases less strongly in unobserved ability if agents have more human capital. Human capital thus reduces the informational rents of high ability types and relaxes the incentive constraints. Since parents do not take the effect of human capital on incentives into account when choosing how much to invest into their children, there is a rationale for education subsidies.

Keywords: human capital, optimal taxation.
JEL: E24, H21, I22, J24.

*University of St.Gallen (Swiss Institute for Empirical Economic Research), CEPR, CFS, IZA. **CNRS (CREST), CEPR, CESifo, IZA. Part of this research has been conducted while Winfried Koeniger was at Queen Mary, University of London; and Julien Prat was at the Institute for Economic Analysis (CSIC).

1 Introduction

The question whether and how much to redistribute between agents with different income and wealth is of paramount importance for most societies. The seminal contribution by Mirrlees (1971) delivered a rigorous framework to answer this normative question, based on the essential trade-off between equality and incentives. In that framework, income is observable but not the ability type of individuals, so that redistribution is constrained by incentive compatibility which ensures that agents truthfully reveal their type.

We build on Mirrlees (1971) and the subsequent literature on optimal taxation by analyzing the problem of optimal redistribution in a model with human capital. Our analysis is motivated by empirical observations for OECD countries in Figure 1 which show how redistribution, measured by a representative marginal tax rate on labor income and bequests, correlates with human capital, measured by the percentage of the population with tertiary education.[1] While the data measures are imperfect and the correlations are not easily interpreted, the data variation in Figure 1 raises the question whether and how human capital affects the optimal amount of redistribution.

We tackle the question with a model of family dynasties in which each generation is altruistic. The working-age generation decides how much to consume, to bequeath in terms of bonds and to invest into human capital of their offspring. Bequests and human capital are observable but the ability type of each generation is not.[2]

We characterize the wedges between the laissez faire and the social optimum for labor supply, bequests in bonds and human capital investment. While the wedges for labor supply and bequests correspond to previous findings in the literature (Farhi and Werning, 2013, Kocherlakota, 2010, and references therein), the wedge for human capital provides novel insights to the best of our knowledge. We find that human capital relaxes the incentive constraints: for standard assumptions about preferences and technology, the disutility of labor decreases less strongly in unobserved ability if agents have more human capital. Human capital thus reduces the informational rents of high-ability agents. Since this effect is not internalized in the laissez faire, there is a rationale for education subsidies.

These findings differ from Findeisen and Sachs (2012) who find an opposite incentive effect of human capital. The reason is a different assumption about

[1] The positive correlation between the inheritance tax and tertiary education only holds for countries that charge a strictly positive inheritance tax. The correlation is similar if we consider the marginal inheritance tax rate for bequests of 100,000 EURO, although fewer countries have a positive tax rate for such an amount. Details on the data sources are provided in appendix A.1.

[2] See Kapicka (2006) for an analysis with unobservable human capital.

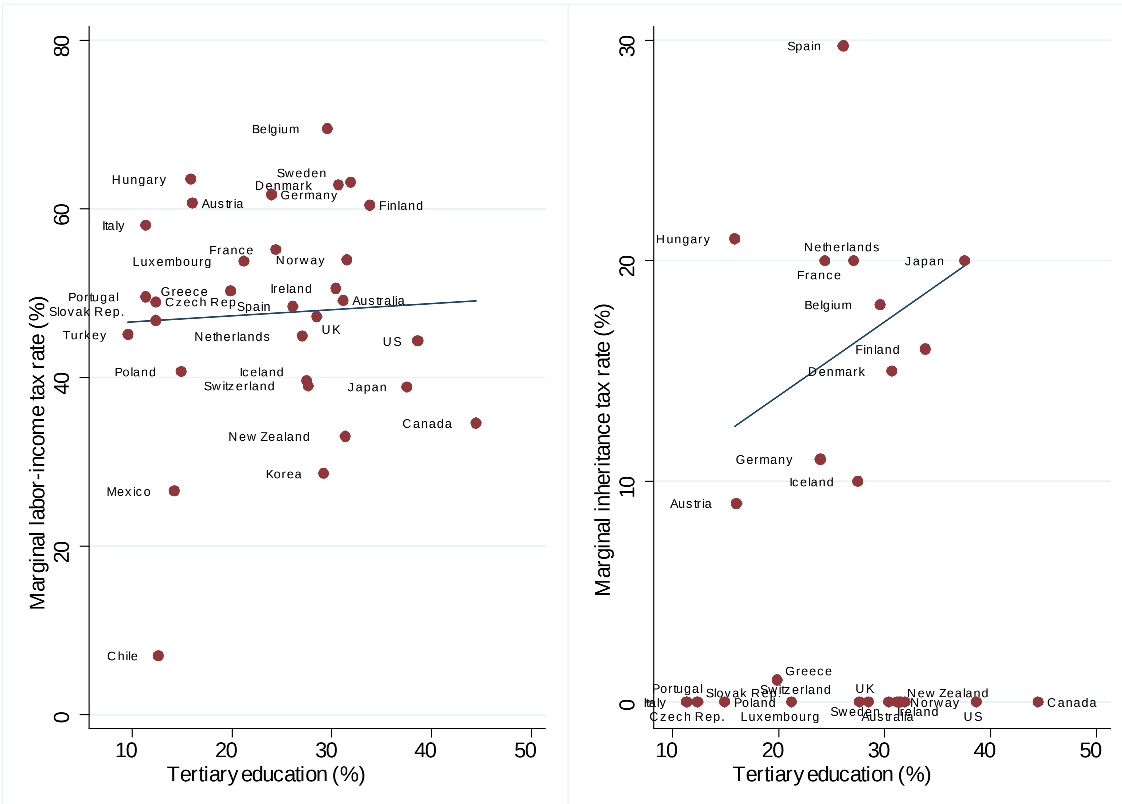

Figure 1: Taxes and human capital in OECD countries in the 2000s. Sources: OECD, CESifo, own compilation. Notes: Tertiary education is measured in % of the population; the tax on labor income is the marginal tax for a worker who earns 133% of the average production wage, following Bovenberg and Jacobs (2005); the marginal inheritance tax is the marginal tax for spouses and children with an inheritance of 250,000 Euro. Details on the data sources are contained in appendix A.1.

how human capital affects productivity. In Findeisen and Sachs (2012), more human capital and higher innate ability both favorably shift the distribution function of labor market productivity but do not enter deterministically as inputs in the production technology. Thus, changes in human capital do not matter with certainty for the amount of labor supply needed to produce a given unit of output. Since Findeisen and Sachs (2012) assume that more human capital reinforces the effect of innate ability on the distribution function of productivity, human capital increases the informational rents of high-ability types. Thus, the incentive compatibility constraint tightens so that it is optimal to tax human capital investment ceteris paribus.

In our model, we assume a standard production technology in which labor productivity depends on human capital and innate ability with an aggregator function that exhibits a constant elasticity of substitution. This technology implies that the disutility of labor effort to produce a given output decreases *less* in innate ability if human capital is higher, for plausible degrees of complementarity between innate ability and human capital. Then, more human capital reduces the effort cost for all agents to produce a given output, and this effect is stronger for agents with low innate ability. It follows that more human capital alleviates the incentive problem so that the planner has a motive to subsidize education.

We find that the incentive effect of human capital is positive if the elasticity of substitution between human capital and innate ability is larger than $1/4$. Human capital thus has a positive incentive effect for the frequently used Cobb-Douglas production function. This finding differs from Stantcheva (2014) who argues that human capital only has a positive incentive effect if the elasticity of substitution between human capital and innate ability is larger than unity. The reason is that we also consider the effect of human capital accumulation on incentives through its effect on labor supply.

The rest of the paper is structured as follows. In Section 2 we describe the model set-up before we solve the planner's problem in Section 3. In Section 4 we derive the optimality conditions in the laissez faire and then characterize the wedges between the laissez faire and the social optimum in Section 5. We conclude in Section 6.

2 Model setup

We take family dynasties as our units of analysis. Each family is composed of parents and children in each generation. The family chooses the labor supply of the parents, bequests and education for the children. Preferences link generations in a time separable fashion. Per period utility is increasing in the family's consumption c and decreasing in labor effort l so that expected lifetime utility $\mathcal{U}$ reads

$$\mathcal{U} \equiv \mathbb{E} \sum_{t=0}^{T} \beta^t \mathbf{U} \left(c_t, l_t \right) \ ,$$

where $\mathbb{E}$ is the expectation operator and β is the discount factor measuring the strength of the altruism towards future generations. The period utility $\mathbf{U} \left(c_t, l_t \right)$ is increasing in consumption c_t and decreasing in labor effort l_t.

As in the seminal paper of Mirrlees (1971), agents have heterogenous ability types $\theta \in [\underline{\theta}; \overline{\theta}]$ which are not observed by the planner. We assume that the planner has information about bequests b_t and human capital h_t. Output y_t of each family is produced using technology

$$y_t = Y \left(h_t, l_t, \theta_t \right),$$

where we assume that $Y \left(\cdot \right)$ is increasing in its arguments and concave. Although output y_t and human capital h_t are observable, this does not reveal labor effort l_t since θ_t is stochastic and not observed.

Human capital of a family in the next period h_{t+1} depends on the expenditure flow for education e_t and family background h_t (the stock of human capital of parents). In the spirit of Ben-Porath (1967), the human capital production function is $h_{t+1}(e_t, h_t)$ where this function is increasing in its arguments and concave.

The timing in the model is as follows. In any given period t, the family learns the parents' type θ_t and chooses to spend e_t on the children's human capital h_{t+1}, to supply parents' labor l_t, to consume c_t and thus bequeath b_{t+1}. We assume that types are drawn from the distribution $F(\theta_{t+1})$ so that the type is i.i.d. across generations. This assumption simplifies the analytic results without changing the main insight that human capital relaxes incentive compatibility constraints.

3 The planner's problem

According to the revelation principle, we can solve the planner's problem by focusing on a direct mechanism so that families truthfully report the type in each generation. Let $\theta^t \equiv \{\theta_0, \theta_1, ..., \theta_t\}$ denote the history of types within a given family. The optimal allocation is history dependent and summarized by $\{c, h, y\} \equiv \left\{ c \left(\theta^t \right), h \left(\theta^t \right), y \left(\theta^t \right) \right\}$. Denoting with $r^t \equiv \{r_0, r_1, ..., r_t\}$ the history of reported types as of date t, r^t equals θ^t if families find it optimal to report truthfully.

The planner discounts the future with the factor q. We assume a small open economy so that there is no feedback between choices of families due to equilibrium price effects. Hence, the planner's problem can be analyzed separately for each family. Furthermore, we assume that planner gives each generation the same weight as the family (see Kocherlakota, 2010, ch. 5 for an analysis that allows for different weights of the planner).

Cost minimization for each family along the equilibrium path requires

$$\min_{\{c,e,y\}} \mathbb{E} \sum_{t=0}^{T} q^t \int_{\Theta} \left(c\left(\theta^t\right) + e\left(\theta^t\right) - y\left(\theta^t\right) \right) dF\left(\theta^t\right) \ ,$$

subject to the incentive compatibility constraint

$$\mathcal{U}\left(\left\{c\left(\theta^t\right), h\left(\theta^t\right), y\left(\theta^t\right)\right\}\right) \geq \mathcal{U}\left(\left\{c\left(r^t\right), h\left(r^t\right), y\left(r^t\right)\right\}\right)$$

for all types and reports that are feasible, and subject to the promise keeping constraint

$$\mathcal{U}\left(\left\{c\left(\theta^t\right), h\left(\theta^t\right), y\left(\theta^t\right)\right\}\right) \geq \omega_0 \, .$$

Note that for a given reported type θ, the choice of h and y implies labor effort l and $\mathcal{U}\left(\left\{c\left(\theta^t\right), h\left(\theta^t\right), y\left(\theta^t\right)\right\}\right)$ denotes the expected lifetime utility evaluated at the equilibrium path on which families truthfully reveal their type.

We write the planner's problem in recursive form. Since the problem is stationary we use a prime "′" to denote the next period. As is standard in the literature (for example, Farhi and Werning, 2013), the recursive formulation requires the use of promised utility V as a state variable. The recursive problem of the planner is:

$$\Gamma(V,h) = \min_{\{c(\theta),y(\theta),h'(\theta),V'(\theta)\}} \left\{ \int_{\Theta} \left[c(\theta) + g(h'(\theta),h) - y(\theta) + q\Gamma\left(V'(\theta), h'(\theta)\right) \right] dF(\theta) \right\}$$

$$\text{s.t. } \omega(\theta) = U\left(c(\theta), y(\theta), \theta, h\right) + \beta V'(\theta), \tag{1}$$

$$V = \int_{\Theta} \omega(\theta)\, dF(\theta), \tag{2}$$

$$\frac{\partial \omega(\theta)}{\partial \theta} = \frac{\partial U(c,y,\theta,h)}{\partial \theta}. \tag{3}$$

where (i) we have used the production function $Y(h,l,\theta)$ to substitute labor effort l in the utility function with a term that depends on y, θ and h, so that $U(c,y,\theta,h) = \mathbf{U}(c,l)$; and (ii) we have inverted the human capital accumulation function $h'(e,h)$ to substitute $e(\theta)$ with $g(h'(\theta),h)$. The first and third constraint are the promise-keeping constraint and incentive-compatibility constraint, respectively. As is common in the literature, the simplified incentive-compatibility constraint obtains if first-order conditions are necessary and sufficient: in the derivation of the incentive compatibility constraint we use that optimal truthful reporting of families requires

$$\left(\frac{\partial U\left(c\left(r\right),y\left(r\right),\theta,h\right)}{\partial c\left(r\right)}\frac{\partial c\left(r\right)}{\partial r}\right)|_{r=\theta}+\left(\frac{\partial U\left(c\left(r\right),y\left(r\right),\theta,h\right)}{\partial y\left(r\right)}\frac{\partial y\left(r\right)}{\partial r}\right)|_{r=\theta}+\beta\frac{\partial V'\left(r\right)}{\partial r}|_{r=\theta}=0.$$

Since general conditions for the validity of the so-called first-order approach do not exist, its validity has to be verified ex post, by checking whether it is indeed not optimal for families not to deviate from the optimum characterized with the first-order conditions.

The Hamiltonian associated with the planner's minimization problem is

$$\begin{aligned}\mathcal{H} = {} & \left[c\left(\omega\left(\theta\right)-\beta V'\left(\theta\right),y\left(\theta\right),\theta,h\right)+g(h'(\theta),h)-y\left(\theta\right)+q\Gamma\left(V'\left(\theta\right),h'(\theta)\right)\right]f\left(\theta\right)\\ & +\lambda\left[V-\omega\left(\theta\right)f\left(\theta\right)\right]\\ & +\mu\left(\theta\right)\left[\partial U\left(c\left(\omega\left(\theta\right)-\beta V'\left(\theta\right),y\left(\theta\right),\theta,h\right),y\left(\theta\right),\theta,h\right)/\partial\theta\right],\end{aligned}$$

where we have substituted consumption using the promise-keeping constraint.

The costate variable satisfies

$$\frac{\partial\mu\left(\theta\right)}{\partial\theta}=-\left[\frac{\partial c\left(\theta\right)}{\partial\omega\left(\theta\right)}-\lambda+\frac{\mu\left(\theta\right)}{f(\theta)}\frac{\partial^{2}U\left(\cdot\right)}{\partial\theta\partial c\left(\theta\right)}\frac{\partial c\left(\theta\right)}{\partial\omega\left(\theta\right)}\right]f(\theta), \quad (4)$$

with the usual boundary conditions $\lim_{\theta\to\underline{\theta}}\mu\left(\theta\right)=0$ and $\lim_{\theta\to\overline{\theta}}\mu\left(\theta\right)=0$.

The first-order conditions read

$$\frac{\partial\mathcal{H}\left(\cdot\right)}{\partial V'\left(\theta\right)} = \left[\frac{\partial c\left(\theta\right)}{\partial V'\left(\theta\right)}+q\frac{\partial\Gamma\left(V'\left(\theta\right),h'\left(\theta\right)\right)}{\partial V'\left(\theta\right)}\right]f(\theta)+\mu\left(\theta\right)\frac{\partial^{2}U\left(\cdot\right)}{\partial\theta\partial c\left(\theta\right)}\frac{\partial c\left(\theta\right)}{\partial V'\left(\theta\right)}=0, \quad (5)$$

$$\frac{\partial\mathcal{H}\left(\cdot\right)}{\partial h'\left(\theta\right)} = \left[\frac{\partial g(h'(\theta),h)}{\partial h'(\theta)}+q\frac{\partial\Gamma\left(V'\left(\theta\right),h'(\theta)\right))}{\partial h'(\theta)}\right]f\left(\theta\right)=0, \quad (6)$$

$$\frac{\partial\mathcal{H}\left(\cdot\right)}{\partial y\left(\theta\right)} = \left[\frac{\partial c\left(\theta\right)}{\partial y\left(\theta\right)}-1\right]f\left(\theta\right)+\mu\left(\theta\right)\left[\frac{\partial^{2}U\left(\cdot\right)}{\partial\theta\partial c\left(\theta\right)}\frac{\partial c\left(\theta\right)}{\partial y\left(\theta\right)}+\frac{\partial^{2}U\left(\cdot\right)}{\partial\theta\partial l\left(\theta\right)}\frac{\partial l\left(\theta\right)}{\partial y\left(\theta\right)}\right]=0. \quad (7)$$

Note that the human capital of the next generation $h'\left(\theta\right)$ in equation (6) affects the incentive-compatibility constraint since it changes the disutility of labor in the future.[3] This becomes explicit in the proof of Proposition 1 in the

[3]Our assumption is that human capital investment affects productivity in the next period (for the next generation) and not in the current period as in Stantcheva (2014). This different assumption seems sensible since Stantcheva analyzes human capital investment of individuals over the life cycle while we focus on human capital investment of parents into their children. While the timing assumption matters for the quantitative analysis, it does not matter for the qualitative insights generated by our analysis.

appendix when inspecting the envelope condition $\partial\Gamma(V,h)/\partial h$, which enters in (6) with a one period lead.[4]

In order to simplify the solution, we make the common assumption that the utility function is separable in consumption and effort:

$$\begin{aligned}\textbf{[A1]} \quad : \quad & \mathbf{U}(c,l) = u(c) - \varphi(l)\,, \\ & u(c) \text{ is increasing in } c \text{ and strictly concave,} \\ & \varphi(l) \text{ is increasing in } l \text{ and strictly convex.}\end{aligned}$$

For the following proposition, we again use the production function $y = Y(h,l,\theta)$ to substitute out labor so that we replace the function for the disutility of labor $\varphi(l)$ with a function $v(y,\theta,h)$ that depends on output, ability and human capital.

Proposition 1 *If [A1] holds, the first-order conditions of the planner problem are*

$$\frac{\partial\mathcal{H}(\cdot)}{\partial V'(\theta)} = \left[-\frac{\beta}{\frac{\partial u(c(\theta))}{\partial c(\theta)}} + q\lambda'(\theta)\right] f(\theta) = 0, \tag{8}$$

$$\begin{aligned}\frac{\partial\mathcal{H}(\cdot)}{\partial h'(\theta)} &= \frac{\partial g(h'(\theta),h)}{\partial h'(\theta)} + q\int_{\Theta}\left(\frac{\frac{\partial v(y'(\theta'),\theta',h'(\theta))}{\partial h'(\theta)}}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}} + \frac{\partial g(h''(\theta'),h'(\theta))}{\partial h'(\theta)}\right) dF(\theta') \qquad (9) \\ &- q\int_{\Theta}\mu'(\theta')\frac{\partial^2 v(y'(\theta'),\theta',h'(\theta))}{\partial\theta'\partial h'(\theta)}d\theta' = 0,\end{aligned}$$

$$\frac{\partial\mathcal{H}(\cdot)}{\partial y(\theta)} = \left[\frac{\frac{\partial v(y(\theta),\theta,h)}{\partial y(\theta)}}{\frac{\partial u(c(\theta))}{\partial c(\theta)}} - 1\right] f(\theta) - \frac{\partial^2 v(y(\theta),\theta,h)}{\partial\theta\partial y(\theta)}\mu(\theta) = 0, \tag{10}$$

with

$$\mu(\theta) = \int_{\underline{\theta}}^{\theta}\left[-\frac{1}{\partial u(c(x))/\partial c(x)} + \lambda\right] dF(x). \tag{11}$$

The proof for this proposition and all other results are contained in appendix A.2.

---

[4] Human capital accumulation would also affect incentive compatibility if it crowded out production of goods through the use of production factors which are rival and in finite supply. Time use for human capital investments into children is an obvious example. We abstract from this since the time effort exerted for human capital accumulation is plausibly as unobservable as is the time effort for production. With two hidden actions, however, we would need to consider joint deviations that make the analysis much less tractable so that we abstract from them for the analysis in this paper.

In order to interpret the first-order conditions for the social optimum, we specify the disutility of labor and the production function of output and human capital as

[A1']: $\varphi(l) = \zeta l^{\alpha}$, with $\zeta > 0$ and $\alpha > 1$,
[A2]: $Y(h, l, \theta) = A(\theta, h)\, l$, with $A(\theta, h) = [\xi \theta^{\chi} + (1 - \xi) h^{\chi}]^{1/\chi}$, $\chi \in (-\infty, 1]$ and $\xi \in (0, 1)$,
[A3]: $h'(e, h) = [\phi e^{\rho} + (1 - \phi) h^{\rho}]^{1/\rho}$ with $\rho \in (-\infty, 1]$ and $\phi \in (0, 1)$,

to show:

Remark 1 *Under assumptions [A1'], [A2] and [A3],*

$$\begin{aligned}
\frac{\partial v(y,\theta,h)}{\partial h} &< 0, \frac{\partial v(y,\theta,h)}{\partial \theta} < 0, \frac{\partial v(y,\theta,h)}{\partial y} > 0, \\
\frac{\partial v(y,\theta,h)}{\partial \theta \partial h} &\geq 0 \text{ iff } \chi \geq -\alpha, \frac{\partial v(y,\theta,h)}{\partial \theta \partial y} < 0, \\
\frac{\partial g(h',h)}{\partial h'} &> 0, \frac{\partial g(h',h)}{\partial h} < 0.
\end{aligned}$$

Before we discuss the meaning of $\chi \geq -\alpha$, let us first mention that:

Corollary 1 *Condition (8) for optimal bequests implies the standard reciprocal Euler equation since*

$$\lambda'(\theta) = \int_{\Theta} \frac{1}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}} dF(\theta').$$

Concerning human capital accumulation, the first two terms of condition (9) equate the marginal cost of human capital investment with the marginal benefit, which consists of the lower disutility of labor to produce a given quantity of output and the reduced cost of accumulating human capital for the next generation. The last term in (9) is the effect of human capital on the incentive compatibility constraint and thus of particular interest. Remark 1 implies that the incentive compatibility constraint

$$\frac{\partial \omega(\theta)}{\partial \theta} = \frac{\partial U(c, y, \theta, h)}{\partial \theta} = -\frac{\partial v(y,\theta,h)}{\partial \theta} > 0.$$

Incentive compatibility thus prevents full insurance against stochastic changes in ability risk and implies that families with higher-ability parents have higher continuation utility.

How does human capital affect incentive compatibility? If human capital is not too complementary to innate ability, $\chi \geq -\alpha$ and thus $v(y,\theta,h)/\partial\theta\partial h > 0$ in condition (9) by Remark 1. Hence, the incentive compatibility constraint is expected to bind less (for given $\mu'(\theta') > 0$) if agents have more human capital. Note that the elasticity of substitution between h and θ is $1/(1-\chi)$ and $\alpha = \varepsilon^{-1} + 1$, where ε is the Frisch elasticity of labor supply. Thus for a plausible value of $\varepsilon = 0.5$, as in Farhi and Werning (2013) or Stantcheva (2014), $\alpha = 3$ and $v(y,\theta,h)/\partial\theta\partial h > 0$ if the elasticity of substitution between human capital and innate ability is larger than $1/4$.

The condition for optimal production (10) is analogous to the optimality condition in the standard Mirrlees problem. Thus, we further comment on it only when we characterize the wedges for which we require the optimality conditions in the laissez faire.

4 The laissez faire

Each family solves the maximization problem

$$\begin{aligned} W(\theta,b,h) &= \max_{\{b',h',l\}} \left\{ \mathbf{U}(c,l) + \beta \int_\Theta W(\theta',b',h')\, dF(\theta') \right\} \\ \text{s.t. } b' &= (1+r)b - c - e + y, \\ y &= Y(h,\theta,l), \\ h' &= h'(e,h) \text{ so that } e = g(h',h), \end{aligned}$$

where b is the bequest.

Proposition 2 *The laissez faire is characterized by the following first-order conditions for bequests, human capital and labor supply:*

$$\begin{aligned} \frac{\partial \mathbf{U}(c,l)}{\partial c} &= \beta(1+r)\mathbb{E}\left[\frac{\partial \mathbf{U}(c',l')}{\partial c'}\right] \\ \frac{\partial g(h',h)}{\partial h'}\frac{\partial \mathbf{U}(c,l)}{\partial c} &= \beta \int_\Theta \left[\frac{\partial y'}{\partial h'}\frac{\partial \mathbf{U}(c',l')}{\partial c'}\right] dF(\theta') \\ &\quad -\beta \int_\Theta \left[\frac{\partial g(h'',h')}{\partial h'}\frac{\partial \mathbf{U}(c',l')}{\partial c'}\right] dF(\theta') \\ -\frac{\partial \mathbf{U}(c,l)}{\partial l} &= \frac{\partial y}{\partial l}\frac{\partial \mathbf{U}(c,l)}{\partial c} \end{aligned}$$

We assume that the problem is concave so that the conditions in Proposition 2, and thus also Proposition 1, are necessary and sufficient. Non-concavities may arise due to the joint decision of labor supply *and* human capital. As is well known from the literature on human capital, concavity can be ensured by restricting parameters to ensure "enough" concavity of the production function in its inputs.

5 The wedges

We combine the results of Propositions 1 and 2 to derive interpretable conditions for the wedges between the choices in the laissez faire and the constrained-efficient allocation of the planner. We start with the following definition.

Definition 1 *The wedges for bequests τ_b, labor supply τ_l and human capital τ_h are*

$$\tau_b(\theta) \equiv 1 - \frac{q}{\beta}\frac{\partial u(c(\theta))/\partial c(\theta)}{\mathbb{E}\left[\partial u(c'(\theta'))/\partial c'(\theta')\right]}, \tag{12}$$

$$\tau_l(\theta) \equiv 1 - \frac{\partial v(y(\theta),\theta,h)/\partial y(\theta)}{\partial u(c(\theta))/\partial c(\theta)}, \tag{13}$$

$$\tau_h(\theta) \equiv \frac{\beta}{\frac{\partial g(h',h)}{\partial h'}}\int_\Theta \left[\frac{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}{\frac{\partial u(c)}{\partial c}}\left(\frac{\partial y'(\theta')}{\partial h'} - \frac{\partial g(h''(\theta'),h')}{\partial h'}\right)\right] dF(\theta') - 1. \tag{14}$$

Wedges are defined as the deviations from the laissez faire. For constrained efficiency a planner thus needs to reduce (increase) bequests, labor supply or human capital, respectively, if the optimality conditions which characterize the social optimum imply $\tau_j > 0$ ($\tau_j < 0$), $j = b, h, l$.

Proposition 3 *Under assumption [A1], the first-order conditions of the planner's problem imply*

$$\tau_b(\theta) = 1 - \frac{1}{\mathbb{E}\left[\frac{1}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}\right]\mathbb{E}\left[\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}\right]}, \tag{15}$$

$$\tau_l(\theta) = -\frac{\partial^2 v(y(\theta),\theta,h)}{\partial\theta\partial y(\theta)}\frac{\mu(\theta)}{f(\theta)}, \tag{16}$$

$$\tau_h(\theta) = A(\theta) + B(\theta) + C(\theta), \tag{17}$$

with

$$
\begin{aligned}
A(\theta) &\equiv \frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_{\Theta} \frac{\partial y'}{\partial h'} \tau_l'(\theta')\, dF(\theta'),\\
B(\theta) &\equiv \frac{1}{\frac{\partial g(h',h)}{\partial h'}} \mathbb{E}\left[\beta \frac{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}{\frac{\partial u(c(\theta))}{\partial c(\theta)}} - q\right] \mathbb{E}\left[\frac{\partial y'(\theta')}{\partial h'} - \frac{\partial g(h''(\theta'),h')}{\partial h'}\right]\\
&\quad + \frac{\beta}{\frac{\partial g(h',h)}{\partial h'}\frac{\partial u(c(\theta))}{\partial c(\theta)}} cov\left(\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}, \frac{\partial y'(\theta')}{\partial h'} - \frac{\partial g(h''(\theta'),h')}{\partial h'}\right),\\
C(\theta) &\equiv -\frac{q}{\frac{\partial g(h',h)}{\partial h'}} cov\left(\frac{1}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}, \frac{\partial v(y'(\theta'),\theta',h')}{\partial h'}\right).
\end{aligned}
$$

By Jensen's inequality, we obtain the result that the wedge for bequests $\tau_b(\theta) > 0$. This wedge implies that the planner wants to reduce bequests to discourage the double deviation that parents leave bequests to their children and children exert little labor effort. The expression for the labor wedge τ_l is standard: by Remark 1 $\partial^2 v(y(\theta),\theta,h)/(\partial\theta\partial y(\theta)) < 0$ and it follows that $\tau_l(\theta) > 0$ if the incentive constraint is binding with $\mu(\theta) > 0$. The intuition is that the additional output generated by labor effort tightens the incentive compatibility constraint, increases the information rents of families and thus allows for less redistribution. Families do not internalize this effect when choosing their optimal labor supply. In order to compare the expression for the labor wedge more easily with the literature, we state:

Corollary 2 *Under assumption **[A1']** and **[A2]***

$$
\frac{\tau_l(\theta)}{1-\tau_l(\theta)} = \alpha \frac{\xi\theta^{\chi}}{A^{\chi}} \frac{\partial u(c(\theta))/\partial c(\theta)}{\theta f(\theta)} \int_{\underline{\theta}}^{\theta} \left[\lambda - \frac{1}{\frac{\partial u(c(x))}{\partial c(x)}}\right] dF(x),
$$

where $\alpha = \varepsilon^{-1} + 1$ and ε denotes the Frisch elasticity of labor supply.

Corollary 2 shows that the labor wedge in our model is analogous to the wedge in Mirrlees (1971).[5]

[5]Note that compared with Mirrlees (1971), the multiplier λ is in the numerator since the shadow price λ is in units of marginal utils and not of public funds of the planner. Furthermore, $\lim_{\theta\to\underline{\theta}} \mu(\theta) = 0$ and $\lim_{\theta\to\overline{\theta}} \mu(\theta) = 0$ imply that

$$
\int_{\underline{\theta}}^{\theta} \left[\lambda - \frac{1}{\frac{\partial u(c(x))}{\partial c(x)}}\right] dF(x) = \int_{\theta}^{\overline{\theta}} \left[\frac{1}{\frac{\partial u(c(x))}{\partial c(x)}} - \lambda\right] dF(x).
$$

The expression for the wedge of human capital $\tau_h(\theta)$ in Proposition 3 is worth discussing further. We decompose this wedge into three terms. The first term $A(\theta)$ relates the human capital wedge $\tau_h(\theta)$ to expectations about the labor wedge $\tau_l'(\theta')$. These expectations are weighed by the marginal product of human capital.[6]

The second term $B(\theta)$ relates τ_h to the wedge for bequests τ_b (see definition (12)). The first term in $B(\theta)$ is positive since

$$\mathbb{E}\left[\beta\frac{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}{\frac{\partial u(c(\theta))}{\partial c(\theta)}}-q\right]>0 \text{ if } \tau_b>0$$

and the marginal product of human capital $\partial y'/\partial h'$ and the cost-reducing effect of human capital on human capital accumulation for the next generation $-\partial g(h''(\theta'),h'(\theta))/\partial h'(\theta)$ are both positive (by Remark 1, $\partial g(h''(\theta'),h'(\theta))/\partial h'(\theta) < 0$). Intuitively, if the planner discourages bequests, constrained efficiency requires also to discourage human capital accumulation since human capital is an alternative way of transferring utility from the current to the future generation. The difference is that the return to human capital depends on ability and is thus risky. This effect is captured by the second term in $B(\theta)$ which depends on the covariance between the return to human capital and the marginal utility of consumption. The covariance is negative if both the return to human capital $\partial y'/\partial h' - \partial g(h''(\theta'),h'(\theta))/\partial h'(\theta)$ and consumption c' of the next generation increase with ability θ' (so that the marginal utility of consumption decreases in θ'). In this case, riskier returns to human capital imply that planner discourages human-capital investment less.

Terms $A(\theta)$ and $B(\theta)$ show that the wedge for human capital accumulation depends on the wedges for bequests and labor supply which is intuitive since human capital alters the marginal product of labor and transfers resources across periods. The last term $C(\theta)$ captures the effect of human capital accumulation on the incentive-compatibility constraint. This effect depends on the covariance between the effect of human capital on the disutility of labor for the next generation and the response of consumption of the next generation to marginal changes in utils (which equals the inverse of the marginal utility of consumption).

Inspecting the covariance, we observe that its sign depends on the sign of $\partial v(y'(\theta'),\theta',h')/\left(\partial h'\partial\theta'\right)$. This cross-derivative captures how the incentive compatibility constraint for the next generation $\partial\omega(\theta')/\partial\theta' = -\partial v(y',\theta',h')/\partial\theta'$ changes

[6] The relationship with the labor wedge in term $A(\theta)$ relates to results in Bovenberg and Jacobs (2005) who show that human capital should be subsidized if taxation of labor income does not condition on human capital. The relationship to Bovenberg and Jacobs (2005) becomes clearer if one analyzes the implementation of the social optimum with taxes and transfers in the decentralized equilibrium.

in human capital. If $\partial v(y'(\theta'), \theta', h') / (\partial h' \partial \theta') > 0$, the incentive compatibility constraint is mitigated by more human capital and informational rents decrease. As shown in the proof of Remark 1, $\partial v(y'(\theta'), \theta', h') / (\partial h' \partial \theta')$ has the sign of the sum $\chi + \alpha$: α captures the effect of human capital on the incentive compatibility constraint due to changes in labor supply and χ captures the effect of human capital on incentive compatibility due to changes in the productivity increase $(\partial A(\theta, h)/\partial \theta) / A(\theta, h)$ resulting from higher ability (in percentage terms).

The effect of human capital on incentives through changes in labor supply unambiguously improves incentives since $\alpha > 0$: more human capital implies that a given unit of output needs to be produced with less labor and the disutility of labor becomes less sensitive in innate ability. Formally, this effect reduces $\partial \omega(\theta') / \partial \theta'$ and thus informational rents.

The effect of human capital on incentives through changes in the productivity increase $(\partial A(\theta, h)/\partial \theta) / A(\theta, h)$ instead depends on the complementarity of human capital and innate ability in production. The effect of human capital on the productivity increase worsens incentives iff $\chi < 0$, in which case human capital and innate ability are more complementary than in the Cobb-Douglas case (the elasticity $1/(1-\chi)$ is less than unity). For this to dominate the incentive-improving effect of human capital through changes in labor supply, human capital and innate ability need to be so complementary that $\chi < -\alpha$.

If $\chi \geq -\alpha$ instead, $\partial v(y'(\theta'), \theta', h') / (\partial h' \partial \theta') > 0$ by Remark 1 and concavity of $u(c)$ implies that the covariance is positive in the term $C(\theta)$ of Proposition 3, as long as consumption $c'(\theta')$ increases in θ'. Thus,

Corollary 3 *Under assumptions **[A1]**, **[A1']** and **[A2]**, $C(\theta) < 0$ if $\chi \geq -\alpha$. The planner then has a motive to increase human capital accumulation in order to relax the incentive compatibility constraint.*

For a plausible Frisch elasticity of labor supply of $\varepsilon = 0.5$, $\alpha = \varepsilon^{-1} + 1 = 3$. Hence, the condition $\chi \geq -\alpha$ implies that the elasticity of substitution between human capital and innate ability in the production function $1/(1-\chi)$ has to be larger than $1/4$. The empirical evidence reviewed by Stantcheva (2014) suggests that this condition for the elasticity of substitution is likely to be satisfied. Our finding that human capital alleviates incentive constraints for empirically plausible parameters differs from Findeisen and Sachs (2012) and Stantcheva (2014) since we consider the effect of human capital accumulation on labor supply.

6 Implementation of the constrained-efficient allocation

TBC

7 Conclusion

We have shown that human capital investment by families is not constrained efficient if the ability of generations in a family is not observable. The wedge for human capital accumulation implied by the solution to the planner's problem depends on the labor wedges for the next generation and the wedge for bequests. We find that, for plausible parameter values, the planner has a motive to increase human capital accumulation in order to relax incentive compatibility constraints. The reason is that more human capital reduces the slope of the disutility of labor in unobservable ability and thus reduces the information rents of high ability types.

A Appendix

A.1 Data sources

This appendix contains information about the data used in Figure 1.

Marginal income tax rate is the marginal tax for a worker who earns 133% of the average production wage including social security contribution rates. We use the average of available observations per country in the period 2000 to 2007. The source is Table I.4 in the OECD Tax database 2013 available at

http://www.oecd.org/tax/tax-policy/Table%20I.4_Mar_2013.xlsx

Marginal inheritance tax rate is the marginal tax for a spouse or child with an inheritance of 250,000 Euro. The data source is the compilation by the CESifo group for 2007 available at

http://www.cesifo-group.de/ifoHome/facts/DICE/Public-Sector/
Public-Finance/Taxes/inheri-tax-rate-07.html

For countries with missing data in 2007 we use information for 2010 available at http://www.cesifo-group.de/ifoHome/facts/DICE/Public-Sector/
Public-Finance/Taxes/inheritance-taxes-key-characteristics-european-union.html

For Poland, the U.S. and Iceland we obtain data from the following sources. The Polish data are available at

http://www.finanse.mf.gov.pl/web/wp/abc-podatkow/asystent-podatnikawe.

For the U.S. we use information on federal taxes available in Figure D, p. 122 in http://www.irs.gov/pub/irs-soi/ninetyestate.pdf.

For Iceland the data are from

http://www.pwc.com/is/is/assets/document/pwc_tax_brochure2013.pdf%20.

Note that the rate we use for Greece applies for real estate; for other assets the tax rate can be higher at 10%.

Tertiary education is the total share of the adult population with tertiary education. The source is the OECD database *Education at a Glance*. We use the average per country in the time period 2000–2007. The data are contained in the publications for the years 2002–2009: in Table A1.3a for 2005–2009, A1.1 for 2004, A2.3 for 2003 and A3.1a for 2002.

A.2 Proofs

Proposition 1

Proof. Under assumption **[A1]** we invert the separable utility function to retrieve consumption as

$$c\left(\omega\left(\theta\right)-\beta V'\left(\theta\right),y\left(\theta\right),\theta,h\right)=u^{-1}\left(\omega\left(\theta\right)-\beta V'\left(\theta\right)+v(y\left(\theta\right),\theta,h)\right).$$

It follows that

$$\begin{aligned}\frac{\partial c\left(\theta\right)}{\partial\omega\left(\theta\right)} &= \frac{1}{\partial u\left(c(\theta)\right)/\partial c(\theta)},\\ \frac{\partial c\left(\theta\right)}{\partial V'\left(\theta\right)} &= -\frac{\beta}{\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)},\\ \frac{\partial c\left(\theta\right)}{\partial y\left(\theta\right)} &= \frac{\partial v(y\left(\theta\right),\theta,h)/\partial y\left(\theta\right)}{\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)},\\ \frac{\partial c\left(\theta\right)}{\partial h} &= \frac{\partial v(y\left(\theta\right),\theta,h)/\partial h}{\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)}.\end{aligned}$$

Condition for V': Since **[A1]** implies $\partial^2 U\left(\cdot\right)/(\partial\theta\partial c)=0$, equation (5) simplifies to

$$\frac{1}{\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)}=\frac{q}{\beta}\frac{\partial\Gamma\left(V'\left(\theta\right),h'(e(\theta),h)\right)}{\partial V'\left(\theta\right)}=\frac{q}{\beta}\lambda'\left(\theta\right),$$

where we have used the envelope condition $\partial\Gamma\left(V,h\right)/\partial V=\lambda$.

Condition for y: Using $\partial^2 U\left(\cdot\right)/(\partial\theta\partial l)=-\frac{\partial y}{\partial l}\frac{\partial^2 v(y,\theta,h)}{\partial\theta\partial y}$ in (7) yields

$$1-\frac{\partial v(y\left(\theta\right),\theta,h)/\partial y\left(\theta\right)}{\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)}=-\frac{\mu\left(\theta\right)}{f(\theta)}\frac{\partial^2 v(y\left(\theta\right),\theta,h)}{\partial\theta\partial y\left(\theta\right)}.$$

Condition for h': The following envelope condition for human capital is obtained after substituting consumption using the promise-keeping constraint, noting that there is a continuum of incentive-compatibility constraints for all θ and that $\frac{\partial^2 U(\cdot)}{\partial c(\theta)\partial\theta} = 0$:

$$\begin{aligned}\frac{\partial\Gamma(V,h)}{\partial h} &= \int_\Theta \left(\frac{\partial c(\theta)}{\partial h} + \frac{\partial g(h'(\theta),h)}{\partial h}\right) dF(\theta) + \int_\Theta \mu(\theta)\frac{\partial^2 U(\cdot)}{\partial\theta\partial h} d\theta \\ &= \int_\Theta \left(\frac{\partial v(y(\theta),\theta,h)/\partial h}{\partial u(c(\theta))/\partial c(\theta)} + \frac{\partial g(h'(\theta),h)}{\partial h}\right) dF(\theta) - \int_\Theta \mu(\theta)\frac{\partial^2 v(y(\theta),\theta,h)}{\partial\theta\partial h} d\theta.\end{aligned}$$

Note the last term which captures the effect of human capital on the incentive compatibility constraint. Note further that for deriving the envelope condition we have inverted $h'(e,h)$ and substituted in $e = g(h',h)$ and we have used that for all θ

$$\begin{aligned}\left(\left(\frac{\partial c(\theta)}{\partial y} - 1\right) f(\theta) + \mu(\theta)\left[\frac{\partial^2 U(\cdot)}{\partial\theta\partial c(\theta)}\frac{\partial c(\theta)}{\partial y(\theta)} + \frac{\partial^2 U(\cdot)}{\partial\theta\partial l(\theta)}\frac{\partial l(\theta)}{\partial y(\theta)}\right]\right)\frac{\partial y(\theta)}{\partial h} &= 0, \\ \left(\frac{\partial g(h'(\theta),h)}{\partial h'(\theta)} + q\frac{\partial\Gamma(V'(\theta),h'(\theta)))}{\partial h'(\theta)}\right)\frac{\partial h'}{\partial h} f(\theta) &= 0\end{aligned}$$

by (6) and (7). The envelope condition for human capital can then be inserted into the optimality condition for human capital (6) so that

$$\begin{aligned}& -q\int_\Theta \left(\frac{\partial v(y'(\theta'),\theta',h')/\partial h'}{\partial u(c'(\theta'))/\partial c'(\theta')} + \frac{\partial g(h''(\theta'),h')}{\partial h'}\right) dF(\theta') \\ & +q\int_\Theta \mu'(\theta')\frac{\partial^2 v(y'(\theta'),\theta',h')}{\partial\theta'\partial h'} d\theta' \\ = \; & \frac{\partial g(h'(\theta),h)}{\partial h'(\theta)}.\end{aligned}$$

For $\frac{\partial^2 U(\cdot)}{\partial c(\theta)\partial\theta} = 0$, equation (4) implies

$$\mu(\theta) = \int_{\underline{\theta}}^{\theta}\left[-\frac{1}{\partial u(c(x))/\partial c(x)} + \lambda\right] dF(x). \tag{18}$$

■

Remark 1

Proof. Inverting the production function $y = Y(h,l,\theta) = A(\theta,h)\,l$, we get $l = y/A(\theta,h)$ with $A(\theta,h) = [\xi\theta^\chi + (1-\xi)h^\chi]^{1/\chi}$ so that

$$\begin{aligned}\frac{\partial v(y,\theta,h)}{\partial y} &= \frac{\partial\varphi\left(\frac{y}{A(\theta,h)}\right)}{\partial\hat{y}} \\ &= \frac{\partial\varphi(l)}{\partial l}\frac{1}{A} \\ &> 0,\end{aligned}$$

$$
\begin{aligned}
\frac{\partial v(y(\theta),\theta,h)}{\partial h} &= \frac{\partial \varphi\left(\frac{y}{A(\theta,h)}\right)}{\partial h} \\
&= -\frac{\partial \varphi(l)}{\partial l}\frac{y}{A^2}\frac{\partial A(\theta,h)}{\partial h} \\
&= -\frac{\partial \varphi(l)}{\partial l} l \frac{\frac{\partial A(\theta,h)}{\partial h}}{A} \\
&= -\frac{\partial \varphi(l)}{\partial l} l (1-\xi) h^{\chi-1} A^{-\chi} \\
&< 0,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial v(y,\theta,h)}{\partial \theta} &= \frac{\partial \varphi\left(\frac{y}{A(\theta,h)}\right)}{\partial \theta} \\
&= -\frac{\partial \varphi(l)}{\partial l}\frac{y}{A^2}\frac{\partial A(\theta,h)}{\partial \theta} \\
&= -\frac{\partial \varphi(l)}{\partial l} l \frac{\frac{\partial A(\theta,h)}{\partial \theta}}{A} \\
&= -\frac{\partial \varphi(l)}{\partial l} l \xi \theta^{\chi-1} A^{-\chi} \\
&< 0,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 v(y,\theta,h)}{\partial \theta \partial y} &= \frac{\partial^2 \varphi\left(\frac{y}{A(\theta,h)}\right)}{\partial \theta \partial y} \\
&= -\frac{\partial^2 \varphi(l)}{\partial l^2}\frac{y}{A^3}\frac{\partial A(\theta,h)}{\partial \theta} - \frac{\partial \varphi(l)}{\partial l}\frac{1}{A^2}\frac{\partial A(\theta,h)}{\partial \theta} \\
&= -\frac{\frac{\partial A(\theta,h)}{\partial \theta}}{A(\theta,h)^2}\frac{\partial \varphi(l)}{\partial l}\left(1 + \frac{l\partial^2 \varphi(l)/\partial l^2}{\partial \varphi(l)/\partial l}\right) \\
&= -\frac{\xi \theta^{\chi-1}}{A^{1+\chi}}\frac{\partial \varphi(l)}{\partial l}\alpha \\
&< 0,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 v(y,\theta,h)}{\partial\theta\partial h} &= \frac{\partial^2 \varphi\left(\frac{y}{A(\theta,h)}\right)}{\partial\theta\partial h} \\
&= \frac{\partial^2\varphi(l)}{\partial l^2}\left(\frac{y}{A^2}\right)^2 \frac{\partial A(\theta,h)}{\partial\theta}\frac{\partial A(\varepsilon,h)}{\partial h} \\
&\quad + \frac{\partial\varphi(l)}{\partial l}\frac{2y}{A^3}\frac{\partial A(\theta,h)}{\partial\theta}\frac{\partial A(\theta,h)}{\partial h} - \frac{\partial\varphi(l)}{\partial l}\frac{y}{A^2}\frac{\partial^2 A(\theta,h)}{\partial\theta\partial h} \\
&= \frac{\partial\varphi(l)}{\partial l}\frac{y}{A^3}\frac{\partial A(\theta,h)}{\partial\theta}\frac{\partial A(\theta,h)}{\partial h}\left(\underbrace{1+\frac{l\partial^2\varphi(l)/\partial l^2}{\partial\varphi(l)/\partial l}}_{\text{Additional term}} + \underbrace{1-\frac{\frac{\partial^2 A(\theta,h)}{\partial\theta\partial h}A(\theta,h)}{\frac{\partial A(\theta,h)}{\partial\theta}\frac{\partial A(\theta,h)}{\partial h}}}_{\text{Stantcheva (2014)}}\right) \\
&= \frac{\partial\varphi(l)}{\partial l}\frac{y}{A}\frac{\xi\theta^{\chi-1}}{A^\chi}\frac{(1-\xi)h^{\chi-1}}{A^\chi}(\alpha+\chi)\ .
\end{aligned}
$$

Thus, $\partial^2 v(y,\theta,h)/(\partial\theta\partial h) > 0$ iff $\chi \geq -\alpha$. Inverting the production function for human capital

$$h'(e,h) = [\phi e^\rho + (1-\phi)h^\rho]^{1/\rho}$$

with $\rho \in (-\infty, 1]$ and $\phi \in (0,1)$, expenditure is given by

$$e = g(h',h) = \left[\frac{1}{\phi}(h')^\rho - \frac{1-\phi}{\phi}h^\rho\right]^{1/\rho}.$$

The expenditure function $g(h',h)$ has constant returns to scale in h' and h and

$$\frac{\partial g(h',h)}{\partial h'} = \frac{1}{\phi}(h')^{\rho-1}g(h',h)^{1-\rho} > 0,$$

$$\frac{\partial g(h',h)}{\partial h} = -\frac{1-\phi}{\phi}h^{\rho-1}g(h',h)^{1-\rho} < 0.$$

■

Corollary 1

Proof. Using equation (4) applying assumption **[A1]**, we have

$$
\begin{aligned}
\int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial\mu(\theta)}{\partial\theta}d\theta &= -\int_{\underline{\theta}}^{\bar{\theta}}\left[\frac{\partial c(\theta)}{\partial\omega(\theta)} - \lambda\right]f(\theta)d\theta \\
&= 0,
\end{aligned}
$$

where the last equality follows from the boundary conditions $\lim_{\theta\to\underline{\theta}}\mu(\theta) = 0$ and $\lim_{\theta\to\bar{\theta}}\mu(\theta) = 0$. Recalling that

$$\frac{\partial c(\theta)}{\partial\omega(\theta)} = \frac{1}{\partial u(c(\theta))/\partial c(\theta)},$$

we find

$$\lambda = \int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{\partial u\left(c(\theta)\right) / \partial c(\theta)} f(\theta) d\theta.$$

Leading this equation one period

$$\lambda'(\theta) = \int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{\partial u\left(c'(\theta')\right) / \partial c'(\theta')} f(\theta') d\theta',$$

equation (8) implies the reciprocal Euler equation. ■

Proposition 2

Proof. *Savings.* The first-order condition for savings reads

$$-\frac{\partial \mathbf{U}\left(c,l\right)}{\partial c} + \beta \int_{\Theta} \frac{\partial W\left(\theta', b', h'\right)}{\partial b'} dF\left(\theta'\right) = 0,$$

which, reinserting the envelope condition

$$\frac{\partial W\left(\theta, b, h\right)}{\partial b} = (1+r) \frac{\partial \mathbf{U}\left(c,l\right)}{\partial c},$$

yields the Euler equation

$$\begin{aligned} \frac{\partial \mathbf{U}\left(c,l\right)}{\partial c} &= \beta(1+r) \frac{\partial \mathbf{U}\left(c',l'\right)}{\partial c'} dF\left(\theta'\right) \\ &= \beta(1+r) \mathbb{E}\left[\frac{\partial \mathbf{U}\left(c',l'\right)}{\partial c'}\right] . \end{aligned}$$

Labor supply. The first-order condition for labor supply reads

$$\frac{\partial \mathbf{U}\left(c,l\right)}{\partial l} + \beta \int_{\Theta} \left[\frac{\partial W\left(\theta', b', h'\right)}{\partial b'} \frac{\partial y}{\partial l}\right] dF\left(\theta'\right) = 0 .$$

The results above imply

$$\beta \int_{\Theta} \left[\frac{\partial W\left(\theta', b', h'\right)}{\partial b'} \frac{\partial y}{\partial l}\right] dF\left(\theta'\right) = \frac{\partial y}{\partial l} \frac{\partial \mathbf{U}\left(c,l\right)}{\partial c}$$

so that the first-order condition for labour supply simplifies to the standard intratemporal condition

$$\frac{\partial \mathbf{U}\left(c,l\right)}{\partial l} + \frac{\partial y}{\partial l} \frac{\partial \mathbf{U}\left(c,l\right)}{\partial c} = 0.$$

Human capital. The first-order condition for human capital accumulation is

$$\beta \int_{\Theta} \left[-\frac{\partial g(h',h)}{\partial h'} \frac{\partial W(\theta',b',h')}{\partial b'} + \frac{\partial W(\theta',b',h')}{\partial h'} \right] dF(\theta') = 0.$$

The envelope condition is

$$\frac{\partial W(\theta',b',h')}{\partial h'} = \frac{\partial y'}{\partial h'} \frac{\partial \mathbf{U}(c',l')}{\partial c'} - \frac{\partial g(h'',h')}{\partial h'} \frac{\partial \mathbf{U}(c',l')}{\partial c'}.$$

Noting that

$$\frac{\partial \mathbf{U}(c,l)}{\partial c} = \beta \int_{\Theta} \frac{\partial W(\theta',b',h')}{\partial b'} dF(\theta')$$

then implies that the first-order condition for human capital simplifies to

$$\begin{aligned} & \frac{\partial g(h',h)}{\partial h'} \frac{\partial \mathbf{U}(c,l)}{\partial c} \\ = \; & \beta \int_{\Theta} \frac{\partial \mathbf{U}(c',l')}{\partial c'} \left[\frac{\partial y'}{\partial h'} - \frac{\partial g(h'',h')}{\partial h'} \right] dF(\theta'). \end{aligned}$$

■

Proposition 3

Proof. The wedge τ_l evaluated at the solution of the planner's problem follows immediately by using the definition for τ_l in the first-order condition (10) of the planner. To derive the analogous expression for τ_b, we recall that $\lambda'(\theta) = \mathbb{E}\left[\frac{1}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}\right]$ and rearrange the definition of τ_b to substitute $\partial u(c(\theta))/\partial c(\theta)$ in condition (8). The wedge for human capital implied by the solution to the planner's problem is obtained by adding τ_h on both sides of condition (9):

$$\begin{aligned} \tau_h(\theta) &= \tau_h(\theta) - \frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_{\Theta} \left(-\frac{\frac{\partial v(y'(\theta'),\theta',h')}{\partial h'}}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}} - \frac{\partial g(h''(\theta'),h')}{\partial h'} \right) dF(\theta') + 1 \\ & \quad - \frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_{\Theta} \mu'(\theta') \frac{\partial^2 v(y'(\theta'),\theta',h')}{\partial \theta' \partial h'} d\theta'. \end{aligned}$$

Substituting in the definition of the wedge $\tau_h(\theta)$ on the right-hand side, we get

$$\begin{aligned} \tau_h(\theta) &= \frac{\beta}{\frac{\partial g(h',h)}{\partial h'}} \int_{\Theta} \left[\frac{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}{\frac{\partial u(c(\theta))}{\partial c(\theta)}} \left(\frac{\partial y'}{\partial h'} - \frac{\partial g(h''(\theta'),h')}{\partial h'} \right) \right] dF(\theta') - 1 \\ & \quad - \frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_{\Theta} \left(-\frac{\frac{\partial v(y'(\theta'),\theta',h')}{\partial h'}}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}} - \frac{\partial g(h''(\theta'),h')}{\partial h'} \right) dF(\theta') + 1 \\ & \quad - \frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_{\Theta} \mu'(\theta') \frac{\partial^2 v(y'(\theta'),\theta',h')}{\partial \theta' \partial h'} d\theta' \end{aligned}$$

which can be rearranged to

$$\begin{aligned}\tau_h(\theta) &= \frac{q}{\frac{\partial g(h',h)}{\partial h'}}\int_\Theta \frac{\partial y'}{\partial h'}\left(1-\frac{\frac{\partial v(y'(\theta'),\theta',h')}{\partial y'(\theta')}}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}\right)dF(\theta') \\ &+\frac{1}{\frac{\partial g(h',h)}{\partial h'}}\int_\Theta\left(\beta\frac{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}}{\frac{\partial u(c(\theta))}{\partial c(\theta)}}-q\right)\left(\frac{\partial y'}{\partial h'}-\frac{\partial g(h''(\theta'),h')}{\partial h'(\theta)}\right)dF(\theta') \\ &-\frac{q}{\frac{\partial g(h',h)}{\partial h'}}\int_\Theta \mu'(\theta')\frac{\partial^2 v(y'(\theta'),\theta',h')}{\partial\theta'\partial h'}d\theta'.\end{aligned}$$

The first term equals $A(\theta)$ using the definition of the labor wedge (13). The second term equals $B(\theta)$ using that $\mathbb{E}(xy) = cov(x,y) + \mathbb{E}(x)\mathbb{E}(y)$.

In the remaining part of the proof, we focus on the last term of $\tau_h(\theta)$ to derive $C(\theta)$. Integrating the integral of the last term by parts,

$$\int_\Theta \mu'(\theta')\frac{\partial^2 v(y'(\theta'),\theta',h')}{\partial\theta'\partial h'}d\theta' = \mu'(\theta')\frac{\partial v(y'(\theta'),\theta',h')}{\partial h'}\Big|_{\underline{\theta'}}^{\overline{\theta'}} - \int_\Theta \frac{\partial\mu'(\theta')}{\partial\theta'}\frac{\partial v(y'(\theta'),\theta',h')}{\partial h'}d\theta'.$$

The first term on the right-hand side is equal to zero because of the boundary conditions for $\mu'(\theta')$. Thus, using (4) imposing assumption **[A1]**, the last term of the wedge $\tau_h(\theta)$ becomes

$$\begin{aligned}&\frac{q}{\frac{\partial g(h',h)}{\partial h'}}\int_\Theta \frac{\partial\mu'(\theta')}{\partial\theta'}\frac{\partial v(y'(\theta'),\theta',h')}{\partial h'}d\theta' \\ =& -\frac{q}{\frac{\partial g(h',h)}{\partial h'}}\int_\Theta\left[\frac{1}{\partial u(c(\theta'))/\partial c(\theta')}-\lambda'(\theta)\right]\frac{\partial v(y'(\theta'),\theta',h')}{\partial h'}f(\theta')d\theta'.\end{aligned}$$

Since by (8),

$$\lambda'(\theta) = \frac{\beta}{q\partial u(c(\theta))/\partial c(\theta)},$$

we get

$$\begin{aligned}&C(\theta) \\ =& -\frac{q}{\frac{\partial g(h',h)}{\partial h'}}\int_\Theta\left[\frac{1}{\partial u(c(\theta'))/\partial c(\theta')}-\frac{\beta}{q\partial u(c(\theta))/\partial c(\theta)}\right]\frac{\partial v(y'(\theta'),\theta',h')}{\partial h'}f(\theta')d\theta'\end{aligned}$$

The integral simplifies since it is equivalent to

$$\begin{aligned}
& \mathbb{E}\left[\frac{1}{\partial u\left(c\left(\theta'\right)\right)/\partial c(\theta')}-\frac{\beta}{q\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)}\right]\mathbb{E}\left[\frac{\partial v(y'\left(\theta'\right),\theta',h')}{\partial h'}\right] \\
& +cov\left(\frac{1}{\partial u\left(c\left(\theta'\right)\right)/\partial c(\theta')}-\frac{\beta}{q\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)},\frac{\partial v(y'\left(\theta'\right),\theta',h')}{\partial h'}\right) \\
= \; & cov\left(\frac{1}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}},\frac{\partial v(y'\left(\theta'\right),\theta',h')}{\partial h'}\right),
\end{aligned}$$

where the second equality follows since the reciprocal Euler equation implies

$$\mathbb{E}\left[\frac{1}{\partial u\left(c\left(\theta'\right)\right)/\partial c(\theta')}-\frac{\beta}{q\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)}\right]=0.$$

This concludes the proof. ■

Corollary 3
Follows immediately from Proposition 3 and Remark 1.

Corollary 2
To compare the labor wedge in our model with the literature, we use definition (13) to derive

$$\begin{aligned}
\frac{\tau_l(\theta)}{1-\tau_l(\theta)} &= \frac{1-\frac{\partial v(y(\theta),\theta,h)/\partial y}{\partial u(c(\theta))/\partial c(\theta)}}{\frac{\partial v(y(\theta),\theta,h)/\partial y}{\partial u(c(\theta))/\partial c(\theta)}} \\
&= \frac{\partial u\left(c(\theta)\right)/\partial c(\theta)}{\partial v(y(\theta),\theta,h)/\partial y(\theta)}\tau_l(\theta).
\end{aligned}$$

Thus, (16) implies that at the solution of the planner's problem,

$$\frac{\tau_l(\theta)}{1-\tau_l(\theta)}=-\frac{\partial u\left(c(\theta)\right)/\partial c(\theta)}{\partial v(y(\theta),\theta,h)/\partial y(\theta)}\frac{\partial^2 v(y\left(\theta\right),\theta,h)}{\partial\theta\partial y\left(\theta\right)}\frac{\mu\left(\theta\right)}{f\left(\theta\right)}.$$

By Remark 1,

$$\begin{aligned}
\frac{\tau_l(\theta)}{1-\tau_l(\theta)} &= \frac{\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)}{\frac{\partial\varphi(l)}{\partial l}\frac{1}{A}}\frac{\xi\theta^{\chi-1}}{A^{1+\chi}}\frac{\partial\varphi\left(l\right)}{\partial l}\alpha\frac{\mu\left(\theta\right)}{f\left(\theta\right)} \\
&= \alpha\frac{\xi\theta^{\chi}}{A^{\chi}}\frac{\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)}{\theta f\left(\theta\right)}\int_{\underline{\theta}}^{\theta}\left[\lambda-\frac{1}{\frac{\partial u(c(x))}{\partial c(x)}}\right]dF(x),
\end{aligned}$$

where we have substituted in $\mu(\theta)$ using (18).

References

[1] Ben-Porath, Yoram (1967): "The Production of Human Capital and the Life Cycle of Earnings," Journal of Political Economy, vol. 75, 352–365.

[2] Bovenberg, A. Lans and Bas Jacobs (2005): "Redistribution and Education Subsidies are Siamese Twins," Journal of Public Economics, vol. 89, 2005–2035.

[3] Farhi, Emmanuel and Iván Werning (2013): "Insurance and Taxation over the Life Cycle," Review of Economic Studies, vol. 80, 596–635.

[4] Findeisen, Sebastian and Dominik Sachs (2012): "Education and Optimal Dynamic Taxation: the Role of Income-Contingent Loans," Working Paper, University of Konstanz.

[5] Kapicka, Marek (2006): "Optimal Income Taxation with Human Capital Accumulation and Limited Record Keeping," Review of Economic Dynamics, vol. 9, 612–639.

[6] Kocherlakota, Narayana R. (2010): *The New Dynamic Public Finance*, Princeton University Press, Princeton, New Jersey.

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

[8] Stantcheva, Stefanie (2014): "Optimal Taxation and Human Capital Policies over the Life Cycle," Manuscript, MIT.