

Koeniger, Winfried; Prat, Julien

**Conference Paper**

## Human Capital and Optimal Redistribution

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik  
- Session: Macro Aspects of Optimal Taxation, No. B07-V3

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Koeniger, Winfried; Prat, Julien (2014) : Human Capital and Optimal Redistribution, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Macro Aspects of Optimal Taxation, No. B07-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100357>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Human Capital and Optimal Redistribution

Winfried Koeniger\* and Julien Prat\*\*

February 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.<sup>1</sup> 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 fully 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.<sup>2</sup>

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 how human capital affects productivity. In Findeisen and Sachs (2012), more

---

<sup>1</sup>The positive correlation between the inheritance tax and tertiary education only holds for countries that charge a strictly positive inheritance tax. Further details on the data are provided in appendix A.1.

<sup>2</sup>See Kapicka (2006) for an analysis with unobservable human capital.

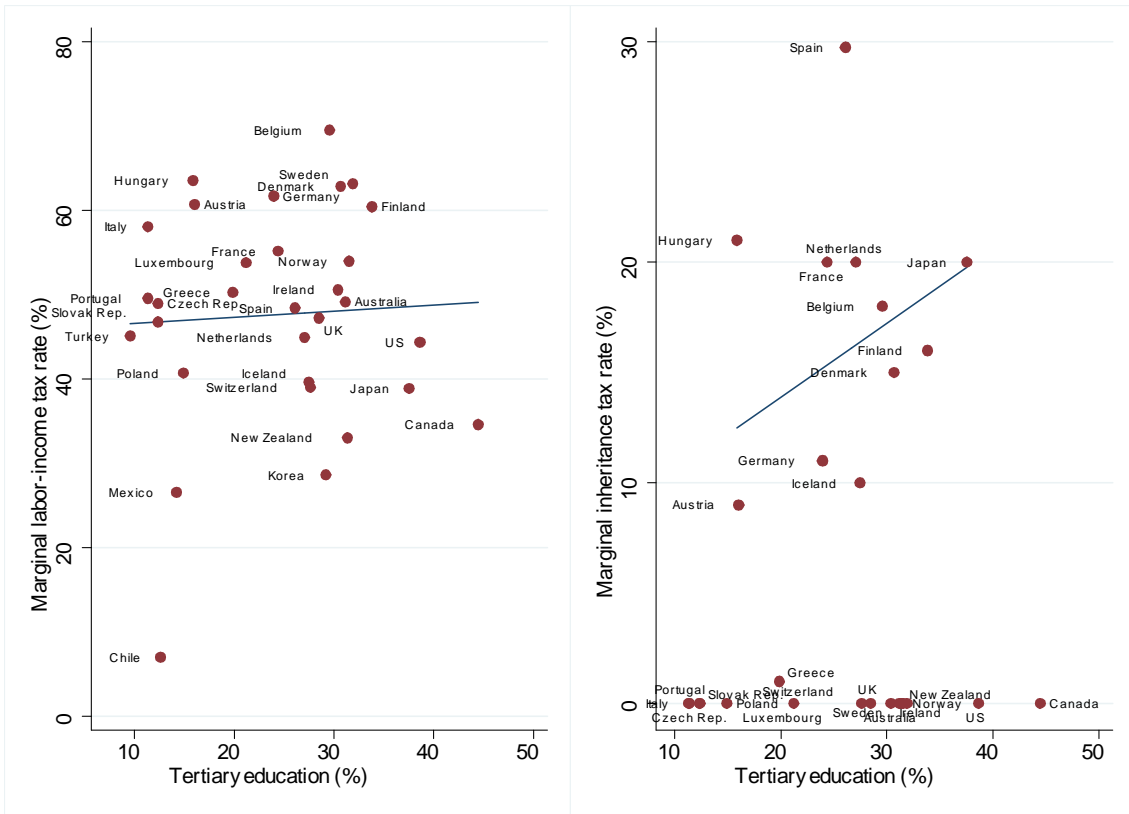


Figure 1: Taxes and human capital in OECD countries in the 2000s. Sources: OECD, CESifo, own compilation. Notes: Tertiary education is measured in % of the population; the tax on labor income is the marginal tax for a worker who earns 133% of the average production wage; the marginal inheritance tax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

human capital and higher innate ability both favorably shift the distribution function of labor market productivity but do not enter directly as inputs in the production technology. Importantly, Findeisen and Sachs (2012) assume that more human capital reinforces the effect of innate ability on the distribution function. Thus, human capital increases the informational rents of high-ability types, tightens the incentive compatibility constraint so that it is optimal to tax human capital investment *ceteris paribus*.

In our model, we assume a standard Cobb-Douglas production technology with human capital, innate ability and labor effort as inputs. This technology implies that the disutility of labor effort to produce a given output decreases *less* in innate ability if human capital is higher. Or, put differently, 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 then follows that more human capital alleviates the incentive problem so that the planner has a motive to subsidize education.

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

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

where  $\mathbb{E}$  is the expectation operator and  $\beta$  is the discount factor measuring the strength of the altruism towards future generations. The period utility  $\mathbf{U}(c_t, l_t)$  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  $\theta \in [\underline{\theta}; \bar{\theta}]$  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

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

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

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  $\theta_t$  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 objective by focusing on a direct mechanism so that families truthfully report the type in each generation. Let  $\theta^t \equiv \{\theta_0, \theta_1, \dots, \theta_t\}$  denote the history of types within a given family. The optimal allocation is history dependent and summarized by  $\{c, h, y\} \equiv \{c(\theta^t), h(\theta^t), y(\theta^t)\}$ . With  $r^t \equiv \{r_0, r_1, \dots, r_t\}$  denoting the history of reported types as of date  $t$  and for optimal truthful reporting  $r^t$  equals  $\theta^t$ .

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

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

subject to the incentive compatibility constraint

$$\mathcal{U}(\{c(\theta^t), h(\theta^t), y(\theta^t)\}) \geq \mathcal{U}(\{c(r_t, \theta^{t-1}), h(r_t, \theta^{t-1}), y(r_t, \theta^{t-1})\})$$

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

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

Note that for a given reported type  $\theta$ , the choice of  $h$  and  $y$  implies labor effort  $l$  and  $\mathcal{U}(\{c(\theta^t), h(\theta^t), y(\theta^t)\})$  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, the recursive formulation requires the use of promised utility  $V$  as a state variable. The recursive problem of the planner is:

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

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

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

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

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  $y$ ,  $\theta$  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. For the derivation of the incentive compatibility constraint, we use that optimal truthful reporting requires

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

The Hamiltonian associated with the planner's minimization problem is

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

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

The costate variable satisfies

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

with the usual boundary conditions  $\lim_{\theta \rightarrow \underline{\theta}} \mu(\theta) = 0$  and  $\lim_{\theta \rightarrow \bar{\theta}} \mu(\theta) = 0$ .

The first-order conditions read

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

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

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

Note that the human capital of the next generation  $h'(\theta)$  affects the incentive-compatibility constraint since it changes the disutility of labor in the future. This becomes explicit in the proof of Proposition 1 in the appendix, when inspecting the envelope condition  $\partial \Gamma(V, h) / \partial h$ .<sup>3</sup>

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

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

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.

---

<sup>3</sup>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.



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

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

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

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

with

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

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

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

$$\mathbf{[A2]:} \quad Y(h, l, \theta) = \theta h^{\kappa_1} l^{\kappa_2}, \quad \kappa_i \geq 0, \quad i = 1, 2, \quad \text{and} \quad \kappa_1 + \kappa_2 \leq 1,$$

$$\mathbf{[A3]:} \quad h' = e^{\gamma_1} h^{\gamma_2}, \quad \gamma_i \geq 0, \quad i = 1, 2, \quad \text{and} \quad \gamma_1 + \gamma_2 \leq 1,$$

to show:<sup>4</sup>

**Remark 2** *Under assumptions [A1], [A2] and [A3],*

$$\begin{aligned} \frac{\partial v(y, \theta, h)}{\partial h} &< 0, \quad \frac{\partial v(y, \theta, h)}{\partial \theta} < 0, \quad \frac{\partial v(y, \theta, h)}{\partial y} > 0, \\ \frac{\partial v(y, \theta, h)}{\partial \theta \partial h} &> 0, \quad \frac{\partial v(y, \theta, h)}{\partial \theta \partial y} < 0, \\ \frac{\partial g(h', h)}{\partial h'} &> 0, \quad \frac{\partial g(h', h)}{\partial h} < 0. \end{aligned}$$

---

<sup>4</sup>The parametric form of the production function matters for the size of the education subsidy in the decentralized equilibrium. The complementarity of human capital to labor and ability, respectively, determines the effect of human capital investment on fiscal revenues and inequality. See Stantcheva (2012).

Let us now interpret the conditions of Proposition 1. Condition (8) implies the reciprocal Euler equation since

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

the gain in consumption due to a marginal increase of utility in the future (see also Kocherlakota, 2010, ch. 5).

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. Since assumption [A1] and Remark 2 imply that the incentive compatibility constraint

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

by  $v(y, \theta, h)/\partial \theta \partial h > 0$  in Remark 2 this incentive compatibility constraint is expected to bind less (for given  $\mu'(\theta') > 0$ ) if agents have more human capital.

Condition (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 3** *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 3, 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 3 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 4** *Under assumption [A1], the wedges for bequests  $\tau_b$ , labor supply  $\tau_l$  and human capital  $\tau_h$  are*

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

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

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

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 5** *The first-order conditions of the planner's problem imply*

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

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

$$\tau_h(\theta) = A(\theta) + B(\theta) + C(\theta), \quad (17)$$

with

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

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 shirk. The expression for the labor wedge  $\tau_l$  is standard, as in the Mirrlees model: by Remark 2  $\partial^2 v(y(\theta), \theta, h) / (\partial \theta \partial y(\theta)) < 0$  and it follows that  $\tau_l(\theta) > 0$  if the incentive constraint is binding with  $\mu(\theta) > 0$ .

The expression for the wedge of human capital  $\tau_h(\theta)$  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  $\tau_l'(\theta')$  (see definition (13)). These expectations are weighed by the marginal product of human capital.<sup>5</sup>

<sup>5</sup>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.

The second term  $B(\theta)$  relates  $\tau_h$  to the wedge for bequests  $\tau_b$  (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 2,  $\partial g(h''(\theta'), h'(\theta)) / \partial h'(\theta) < 0$ ). Intuitively, if the planner discourages bequests, constrained efficiency requires also to discourage human capital accumulation since human capital is alternative way of transferring utility from the current to the future generation. The difference is that the return to human capital depends on ability. 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  $\theta'$  (so that the marginal utility of consumption decreases in  $\theta'$ ).

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). Since utility is concave in consumption and  $\partial v(y'(\theta'), \theta', h') / (\partial h' \partial \theta') > 0$  by Remark 2, the covariance is positive if consumption  $c'(\theta')$  increases in  $\theta$ . Thus,  $C(\theta) < 0$  and the planner has a motive to increase human capital accumulation in order to relax the incentive compatibility constraint. This result differs from Findeisen and Sachs (2012) due to the modeling of the returns to human capital, as explained in the introduction.

If we specify the disutility of labor as in Farhi and Werning (2013),  $\varphi(l) = \zeta l^\alpha$ , with  $\zeta > 0$  and  $\alpha > 1$ , we obtain an expression for the labor wedge which can be compared more easily with the literature.

**Corollary 6** *If  $\varphi(l) = \zeta l^\alpha$ , with  $\zeta > 0$  and  $\alpha > 1$ ,*

$$\frac{\tau_l(\theta)}{1 - \tau_l(\theta)} = \frac{\alpha}{\kappa_2} \frac{\partial u(c(\theta)) / \partial c(\theta)}{\theta f'(\theta)} \int_{\underline{\theta}}^{\theta} \left[ \lambda - \frac{1}{\frac{\partial u(c(x))}{\partial c(x)}} \right] dF(x),$$

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

Corollary 6 shows that the labor wedge in our model is analogous to the wedge in Mirrlees (1971).<sup>6</sup>

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

<sup>6</sup>Note that compared with Mirrlees (1971), the multiplier  $\lambda$  is in the numerator since the shadow price  $\lambda$  is in units of marginal utils and not of public funds of the planner. Furthermore,  $\lim_{\theta \rightarrow \underline{\theta}} \mu(\theta) = 0$  and  $\lim_{\theta \rightarrow \bar{\theta}} \mu(\theta) = 0$  imply that

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

*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/asystem-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](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

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

It follows that

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

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

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

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

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

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

*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  $\theta$  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  $\theta$

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

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

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



■

**Remark 2**

**Proof.** Inverting the production function  $Y(h, l, \theta) = \theta h^{\kappa_1} l^{\kappa_2}$ , we get  $l = (y/(\theta h^{\kappa_1}))^{1/\kappa_2}$  so that

$$\begin{aligned} \frac{\partial v(y, \theta, h)}{\partial y} &= \frac{\partial \varphi(l)}{\partial l} \frac{1}{\kappa_2} \left( \frac{y}{\theta h^{\kappa_1}} \right)^{\frac{1}{\kappa_2}-1} \frac{1}{\theta h^{\kappa_1}} \\ &= \frac{\partial \varphi(l)}{\partial l} \frac{l}{\kappa_2 y} \\ &> 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial v(y(\theta), \theta, h)}{\partial h} &= \frac{\partial \varphi(l)}{\partial l} \frac{1}{\kappa_2} \left( \frac{y}{\theta h^{\kappa_1}} \right)^{\frac{1}{\kappa_2}-1} \left( -\kappa_1 \frac{y}{\theta h^{\kappa_1+1}} \right) \\ &= -\frac{\partial \varphi(l)}{\partial l} \frac{\kappa_1 l}{\kappa_2 h} \\ &< 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial v(y, \theta, h)}{\partial \theta} &= -\frac{\partial \varphi(l)}{\partial l} \frac{1}{\kappa_2} \frac{l}{\theta} \\ &< 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 v(y, \theta, h)}{\partial \theta \partial y} &= \frac{\partial \left( \frac{\partial \varphi(l)}{\partial l} \frac{l}{\kappa_2 y} \right)}{\partial \theta} \\ &= \frac{1}{\kappa_2 y} \left( \frac{\partial^2 \varphi(l)}{\partial l^2} l + \frac{\partial \varphi(l)}{\partial l} \right) \frac{\partial l}{\partial \theta} \\ &= -\frac{1}{\kappa_2 y} \left( \frac{\partial^2 \varphi(l)}{\partial l^2} l + \frac{\partial \varphi(l)}{\partial l} \right) \frac{1}{\kappa_2} \left( \frac{y}{\theta h^{\kappa_1}} \right)^{\frac{1}{\kappa_2}-1} \frac{y}{\theta^2 h^{\kappa_1}} \\ &= -\frac{1}{(\kappa_2)^2 \theta y} \frac{\partial \varphi(l)}{\partial l} l \left( \frac{\partial^2 \varphi(l)}{\partial l^2} \frac{l}{\frac{\partial \varphi(l)}{\partial l}} + 1 \right) \\ &< 0, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 v(y, \theta, h)}{\partial \theta \partial h} &= \frac{\partial \left( -\frac{\partial \varphi(l)}{\partial l} \frac{\kappa_1 l}{\kappa_2 h} \right)}{\partial \theta} \\ &= -\frac{\partial \varphi(l)}{\partial l} \frac{\kappa_1}{\kappa_2} \frac{1}{h} \frac{\partial l}{\partial \theta} - \frac{\partial^2 \varphi(l)}{\partial l^2} \frac{\kappa_1 l}{\kappa_2 h} \frac{\partial l}{\partial \theta} \\ &= \frac{\kappa_1}{(\kappa_2)^2 h \theta} \frac{\partial \varphi(l)}{\partial l} l \left( \frac{\partial^2 \varphi(l)}{\partial l^2} \frac{l}{\frac{\partial \varphi(l)}{\partial l}} + 1 \right) \\ &> 0. \end{aligned}$$

Inverting the production function for human capital  $h' = e^{\gamma_1} h^{\gamma_2}$ , we get  $e = g(h', h) = (h'/h^{\gamma_2})^{\frac{1}{\gamma_1}}$  so that

$$\frac{\partial g(h', h)}{\partial h'} = \frac{1}{\gamma_1} \frac{g(h', h)}{h'} > 0,$$

$$\frac{\partial g(h', h)}{\partial h} = -\frac{\gamma_2}{\gamma_1} \frac{g(h', h)}{h} < 0.$$

■

### Proposition 3

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

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

which, reinserting the envelope condition

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

yields the Euler equation

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

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

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

The results above imply

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

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

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

*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 5

**Proof.** The wedge  $\tau_l$  evaluated at the solution of the planner's problem follows immediately by using the definition for  $\tau_l$  in the first-order condition (10) of the planner. To derive the analogous expression for  $\tau_b$ , we recall that  $\lambda'(\theta) = \mathbb{E} \left[ \frac{1}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}} \right]$  and rearrange the definition of  $\tau_b$  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  $\tau_h$  on both sides of condition (9):

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

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

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

which can be rearranged to

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

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

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,

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

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

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

Since by (8),

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

we get

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

The integral simplifies since it is equivalent to

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

where the second equality follows since the reciprocal Euler equation implies

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

This concludes the proof. ■

**Corollary 6**

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

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

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

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

By Remark 2 and for  $\varphi(l) = \zeta l^\alpha$ ,

$$\begin{aligned} \frac{\tau_l(\theta)}{1 - \tau_l(\theta)} &= \frac{\partial u(c(\theta)) / \partial c(\theta)}{\frac{\partial \varphi(l)}{\partial l} \frac{l}{\kappa_2 y}} \frac{1}{(\kappa_2)^2 \theta y} \frac{\partial \varphi(l)}{\partial l} l \left( \frac{\partial^2 \varphi(l)}{\partial l^2} \frac{l}{\frac{\partial \varphi(l)}{\partial l}} + 1 \right) \frac{\mu(\theta)}{f(\theta)} \\ &= \frac{\alpha}{\kappa_2} \frac{\partial u(c(\theta)) / \partial c(\theta)}{\theta f(\theta)} \int_{\underline{\theta}}^{\theta} \left[ \lambda - \frac{1}{\frac{\partial u(c(x))}{\partial c(x)}} \right] dF(x), \end{aligned}$$

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 (2012): “Optimal Taxation and Human Capital Policies over the Lifecycle,” Slides, MIT.