Human Capital and Optimal Redistribution

Winfried Koeniger* and Julien Prat**

this version: August 2014, first version: February 2014

Preliminary and incomplete
(Please do not circulate)

Abstract

We characterize optimal redistribution in a dynastic family model with human capital accumulation. We show that, under common assumptions about preferences and technology, human capital reduces the informational rents of high ability types and relaxes the incentive constraints. Since parents do not take this externality into account when choosing how much to invest into their children, the model provides 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, Golosov et al. 2011, Kapička, 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 inheritance tax and tertiary education are positively correlated 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.

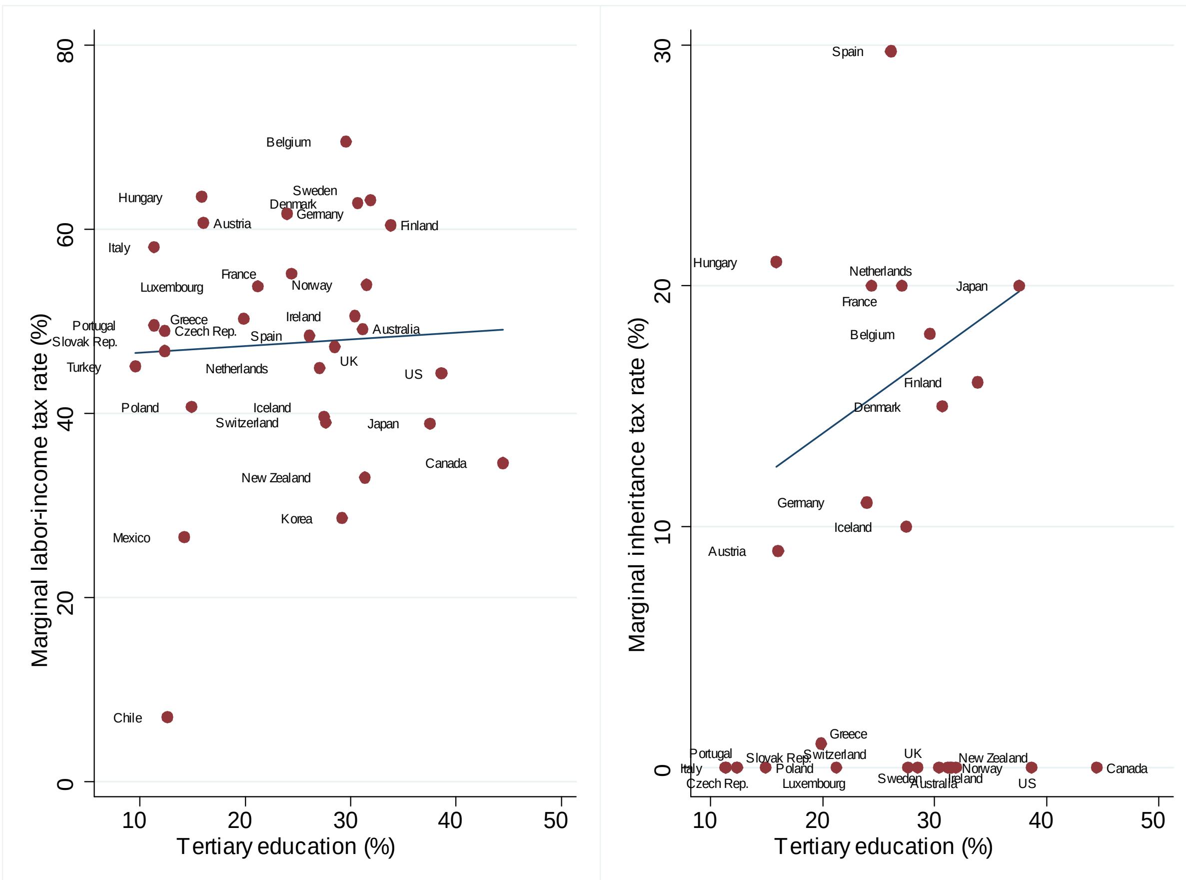

Figure 1: Taxes and human capital in OECD countries in the 2000s. Sources: OECD, CESifo, own compilation. Notes: Tertiary education is measured in % of the population; the tax on labor income is the marginal tax for a worker who earns 133% of the average production wage, following Bovenberg and Jacobs (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 consider the effect of human capital accumulation on incentives through its effect on labor supply while Stantcheva (2014) holds labor supply constant when deriving the effect of human capital on the incentive compatibility constraint in the planner's problem. Since the planner chooses output and cannot directly choose the labor supply of agents, we find that optimal redistribution has to take into account how human capital affects incentives through changes in labor supply.

Our paper is related to the two large literatures on human capital and on optimal taxation reviewed in Stantcheva (2014). Let us only mention a few related and recent papers. Gelber and Weinzierl (2014) analyze optimal taxation if the ability of future generations depends on the resources of the current generation. This is modelled by the letting the probability of types directly depend on disposable income. Our model shares the feature that current resources may impact the earnings capacity of future generations but lets generations choose the amount of resources allocated to human capital accumulation. This allows us to analyze whether that choice is constrained efficient. The papers by Krueger and Ludwig (2013) and Lee and Seshadri (2014) do not use the Mirrlees approach to analyze the effect of redistribution in models with human capital accumulation. Following the Ramsey approach, they specify parametric tax schedules and then analyze the welfare effects of changes in the parameters.

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

2 The planner's problem

2.1 The environment

Family dynasties are the decision units in our 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

$$\mathcal{U} \equiv \mathbb{E}_0 \left[\sum_{t=0}^{T} \beta^t \mathbf{U}(c_t, l_t) \right] ,$$

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 $\mathbf{U}(c_t, l_t)$ is increasing in consumption c_t and decreasing in labor effort l_t. We simplify matters further by making the common assumption that the utility function is separable in consumption and effort

$$\begin{aligned} \textbf{[A1]} \quad : \quad & \mathbf{U}(c,l) = u(c) - \mathbf{v}(l) \, , \\ & u(c) \in \mathcal{C}^2\left(\mathbb{R}^+\right) \text{ is increasing in } c \text{ and strictly concave,} \\ & \mathbf{v}(l) \in \mathcal{C}^2\left(\mathbb{R}^+\right) \text{ is increasing in } l \text{ and strictly convex.} \end{aligned}$$

As in the seminal paper of Mirrlees (1971), agents' abilities $\theta \in \Theta = [\underline{\theta}; \overline{\theta}]$ are heterogenous and cannot be observed by the planner. By contrast, both bequests b_t and human capital h_t are public knowledge. The technology of production $y_t = Y(h_t, l_t, \theta_t)$ is increasing in its arguments and concave. Although output y_t is also observable, the planner cannot use it to infer actual labor supply l_t because the realization of θ_t is hidden.

Human capital in the next period h_{t+1} depends on the expenditure flow for education e_t and on the family background, which can be summarized by the stock of human capital of parents h_t.[3] In the spirit of Ben-Porath (1967), we define a human capital production function $h_{t+1}(e_t, h_t)$ that is increasing in its arguments and concave.[4]

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 abilities are uncorrelated across generations with types being drawn at the beginning of each period from the stationary distribution $F(\theta_{t+1})$. This assumption simplifies the analytic results without changing the main insight that human capital relaxes the incentive compatibility constraints. We thus delegate the results for the model with persistent income shocks to appendix A.3.

2.2 The optimal allocation

According to the revelation principle, we can solve the planner's problem by focusing on a direct mechanism such that families truthfully report their types in each generation. Let $\theta^t \equiv \{\theta_0, \theta_1, ..., \theta_t\}$ denote the history of types within a given family. We do not impose any arbitrary restrictions on the allocation. In particular, we do not rule out history dependent allocations 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 behave truthfully.

The planner discounts the future with the factor q. As Farhi and Werning (2013), we abstract from feedbacks between choices of families due to equilibrium price effects so that the allocation problem can be analyzed separately for each family. Cost minimization for each family along the equilibrium path requires

$$\min_{\{c,e,y\}} \mathbb{E}_0 \left[\sum_{t=0}^{T} q^t \int_{\Theta} \left[c\left(\theta^t\right) + e\left(\theta^t\right) - y\left(\theta^t\right)\right] dF\left(\theta^t\right)\right],$$

subject to the incentive compatibility constraint

$$\mathcal{U}\left(\left\{c\left(\theta^t\right), h'\left(\theta^t\right), y\left(\theta^t\right)\right\}\right) \geq \mathcal{U}\left(\left\{c\left(r^t\right), h'\left(r^t\right), y\left(r^t\right)\right\}\right) \tag{1}$$

[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.

[4]We abstract from time use for human capital investments into children because 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.

for all types and reports which are feasible; and subject to the promise keeping constraint

$$\mathcal{U}\left(\left\{c\left(\theta^t\right), h'\left(\theta^t\right), y\left(\theta^t\right)\right\}\right) \geq \omega_0$$

which ensures that the expected lifetime utility $\mathcal{U}(\cdot)$ of truthful families is at least as high as the exogenously given constraint ω_0. Note that for a given reported type θ, observing h and y pins down effort l. Thus we can use the production function $Y(h, l, \theta)$ to substitute l in the utility function and write $U(c, y, \theta, h) = \mathbf{U}(c, l)$, or with assumption **[A1]**, $v(y, \theta, h) = \mathbf{v}(l)$.

Instead of directly looking for a solution of the allocation problem, we apply two common modifications that simplify the problem considerably. First, we write the planner's problem in recursive form.[5] As shown by Abreu et al. (1990), if ability θ follows an i.i.d. process, we do not need to condition allocations on the entire history of reports but only on the current realization of the promised utility V. Second, we replace the general incentive constraints (1) by an envelope condition that is valid on the equilibrium path on which families truthfully reveal their type. The recursive form of this relaxed planning problem reads

$$\begin{aligned}
\Gamma(V, h, t) &= \min_{\{c, y, h', V'\}} \left\{ \int_\Theta \left[c(\theta) + g(h'(\theta), h) - y(\theta) + q\Gamma\left(V'(\theta), h'(\theta), t+1\right)\right] dF(\theta) \right\} \\
\text{s.t. } \omega(\theta) &= U\left(c(\theta), y(\theta), \theta, h\right) + \beta V'(\theta), && (2) \\
V &= \int_\Theta \omega(\theta)\, dF(\theta), && (3) \\
\frac{\partial \omega(\theta)}{\partial \theta} &= \frac{\partial U(c, y, \theta, h)}{\partial \theta}, && (4)
\end{aligned}$$

where we have inverted the human capital accumulation function $h'(e, h)$ to substitute $e(\theta)$ with $g(h'(\theta), h)$. The first constraint defines the continuation value $\omega(\theta)$ as the sum of the current and next period promised utilities, $U(\cdot)$ and $V'(\theta)$ respectively. Equation (3) is the promise-keeping constraint since it ensures that the expected value of the continuation utility is equal to the promised value V. The last equation is the local incentive-compatibility constraint captured by the envelope condition which is derived assuming that the first-order condition for truthful reporting is satisfied.[6] Since general conditions for the

[5] We only keep a time index for the value function, otherwise drop the indexes and use a prime $'$ to denote the next period. This simplifies notation although it is not fully precise without an infinite horizon.

[6] Totally differentiating the continuation value of a truthful family yields

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

The sum of the last two terms on the right hand side equals zero if the first-order condition for truthful reporting is satisfied, so that we obtain condition (4).

validity of the first-order approach are not available for our problem, incentive compatibility has to be verified ex post by checking whether families do not deviate from the solution of the relaxed problem.[7]

2.3 Optimality conditions

The Hamiltonian associated with the planner's dual minimization problem reads

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

where we have substituted consumption using the definition of $\omega(\theta)$ in (2). In the first best without information asymmetries, $\mu(\theta) = 0$ for all θ and agents are fully insured against transitory changes in ability: consumption remains constant across families and is therefore separated from production. With information asymmetries instead, the planner faces an insurance-incentive trade-off whose optimal balance is determined by the following conditions.

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

$$\frac{\partial \mathcal{H}(\cdot)}{\partial V'(\theta)} = \left[-\frac{\beta}{\frac{\partial u(c(\theta))}{\partial c(\theta)}} + q\lambda'(\theta)\right] f(\theta) = 0, \tag{5}$$

$$\begin{aligned}\frac{\partial \mathcal{H}(\cdot)}{\partial h'(\theta)} &= \frac{\partial g(h'(\theta), h)}{\partial h'(\theta)} + q\int_{\Theta}\left(\frac{\frac{\partial v(y'(\theta'),\theta',h'(\theta))}{\partial h'(\theta)}}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}} + \frac{\partial g(h''(\theta'), h'(\theta))}{\partial h'(\theta)}\right) dF(\theta') \qquad (6) \\ &\quad - q\int_{\Theta}\mu'(\theta')\frac{\partial^2 v(y'(\theta'),\theta',h'(\theta))}{\partial\theta'\partial h'(\theta)}d\theta' = 0,\end{aligned}$$

$$\frac{\partial \mathcal{H}(\cdot)}{\partial y(\theta)} = \left[\frac{\frac{\partial v(y(\theta),\theta,h)}{\partial y(\theta)}}{\frac{\partial u(c(\theta))}{\partial c(\theta)}} - 1\right] f(\theta) - \frac{\partial^2 v(y(\theta),\theta,h)}{\partial\theta\partial y(\theta)}\mu(\theta) = 0, \tag{7}$$

with

$$\mu(\theta) = \int_{\underline{\theta}}^{\theta}\left[\lambda - \frac{1}{\partial u(c(x))/\partial c(x)}\right] dF(x), \textit{ and } \lim_{\theta\to\underline{\theta}}\mu(\theta) = \lim_{\theta\to\bar{\theta}}\mu(\theta) = 0. \tag{8}$$

Consumption.—Equation (5) implies that the reciprocal Euler equation continues to hold when human capital is introduced. To see why, note that evaluating the law of motion (8) of the costate variable at the upper bound of the ability

[7] See the discussion in Kapička (2013) or Fahri and Werning (2013).

distribution yields

$$\lambda - \int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial c(\theta)}{\partial \omega(\theta)} dF(\theta) = \mu(\overline{\theta}) = 0.$$

Replacing $\partial c(\theta)/\partial\omega(\theta) = [\partial u(c(\theta))/\partial c(\theta)]^{-1}$ and leading this equation one period ahead, we find that $\lambda'(\theta) = \int_{\underline{\theta}}^{\overline{\theta}} [\partial u(c'(\theta'))/\partial c'(\theta')]^{-1} dF(\theta')$. Thus the inverse of the marginal utility of consumption follows a martingale:

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

Output.—The condition for optimal production (7) is analogous to the optimality condition in the standard Mirrlees problem. Thus, we further comment on it only when we characterize the wedges in the next section.

Human capital.—Turning our attention to education, let us repeat the optimality condition

$$\frac{\partial g(h',h)}{\partial h'} = -q\int_{\Theta}\left(\frac{\frac{\partial v(y'(\theta'),\theta',h'(\theta))}{\partial h'(\theta)}}{\frac{\partial u(c'(\theta'))}{\partial c'(\theta')}} + \frac{\partial g(h'',h')}{\partial h'}\right) dF(\theta') + q\int_{\Theta}\mu'(\theta')\frac{\partial^2 v(y',\theta',h')}{\partial\theta'\partial h'}d\theta'. \tag{9}$$

The condition equates the marginal cost of human capital investment with the marginal benefit. The marginal benefit has three components. Firstly, human capital lowers the disutility of labor to produce a given quantity of output. This allows the planner to spend less on consumption and still provide the family with the same continuation value. As shown in the proof of Proposition 1 in appendix A.2, this benefit for the planner is captured by $-\frac{\partial v(y'(\theta'),\theta',h'(\theta))}{\partial h'(\theta)}/\frac{\partial u(c'(\theta'))}{\partial c'(\theta')} > 0$. Secondly, more human capital reduces the cost of accumulating human capital for the next generation, $\partial g(h'',h')/\partial h' < 0$.[8] Thirdly, human capital changes the incentive compatibility constraint (the second integral on the right hand side). This term is key for our analysis so that we elaborate on it.

In the absence of informational frictions, families are perfectly insured against transitory shocks so that $\partial\omega(\theta)/\partial\theta = 0$. With hidden types, information revelation is profitable solely if

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

where the inequality follows under the assumption that higher ability reduces the disutility of labor so that $\partial v(\cdot)/\partial\theta < 0$. Incentive compatibility thus prevents

[8] If education costs are independent of family background, $\partial g(h'',h')/\partial h' = 0$.

full insurance: families with more able parents enjoy higher continuation utilities. An increase in the slope $|\partial v(\cdot)/\partial\theta|$ of the disutility term widens the gap separating the optimal allocation from the first best. Hence, the cross-derivative $\partial^2 v(\cdot)/(\partial\theta\partial h)$ measures the effect that human capital has on the incentive compatibility constraint: if $\partial^2 v(\cdot)/(\partial\theta\partial h) > 0$, informational rents decrease which mitigates the incentive problem.

These gains can be translated into consumption units through multiplication by the costate variable $\mu'(\theta')$ which measures the marginal cost of violating the incentive constraint. The resulting products in (9) are integrated over all potential realizations of θ' because neither the planner nor the family know the value of θ' when the human-capital investment decision is made.[9]

The sign of the cross derivative $\partial^2 v(\cdot)/(\partial\theta\partial h)$ is determined by: (i) the Frisch elasticity of labor supply and, (ii) the degree of complementarity between human capital and ability. Both are captured by a single parameter if we assume that the disutility of labor and the production function for output have the following functional forms.

Corollary 1 *Assume that*

[A1']: $\mathrm{U}(c,l) = u(c) - \mathbf{v}(l)$, ***where*** $\mathbf{v}(l) = \zeta l^{\alpha}$, ***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*** $\xi \in (0,1)$.

Then $\partial^2 v(y,\theta,h)/(\partial\theta\partial h) \geq 0$ ***if and only if*** $\chi \geq -\alpha$.

The Frisch elasticity of labor supply is measured by the parameter α and the degree of complementarity by the parameter χ. If the production function is Cobb Douglas, $\chi = 0$. Hence, negative χ imply more complementarity between ability and human capital than in the Cobb-Douglas case. Corollary 1 shows that informational rents are *decreasing* in human capital when the sign of $\chi + \alpha$ is positive, that is when the Frisch elasticity of labor supply α is greater than the degree of complementarity χ (implied by a possibly negative χ) between ability and human capital.

This result is most easily explained using Figure 2 which plots the disutility $v(\cdot)$ as a function of supplied labor. The vertical lines display the values of $l = \hat{y}/A(\theta,h)$ resulting from different combinations of θ and h for a fixed level of output $\hat{y}$. The differences reported on the vertical axis measure the values of $\Delta v(h_i) \equiv v(\hat{y},\theta_2,h_i) - v(\hat{y},\theta_1,h_i) < 0$, with $\theta_2 > \theta_1$. The derivative of the incentive constraint is given by $\partial v(\hat{y},\theta_2,h_i)/\partial\theta = \lim_{\theta_1\to\theta_2} \Delta v(h_i)/(\theta_2-\theta_1)$ and the cross derivative $\partial^2 v(\hat{y},\theta_2,h_2)/(\partial\theta\partial h)$ is of the same sign as $\Delta v(h_2) - \Delta v(h_1)$ when both $\theta_1 \to \theta_2$ and $h_1 \to h_2$. Note that $\Delta v(h_1) < \Delta v(h_2) < 0$ in Figure 2.

[9] Remember that the costate variable $\mu'(\theta')$ contains weights of the density function $f(\theta')$.

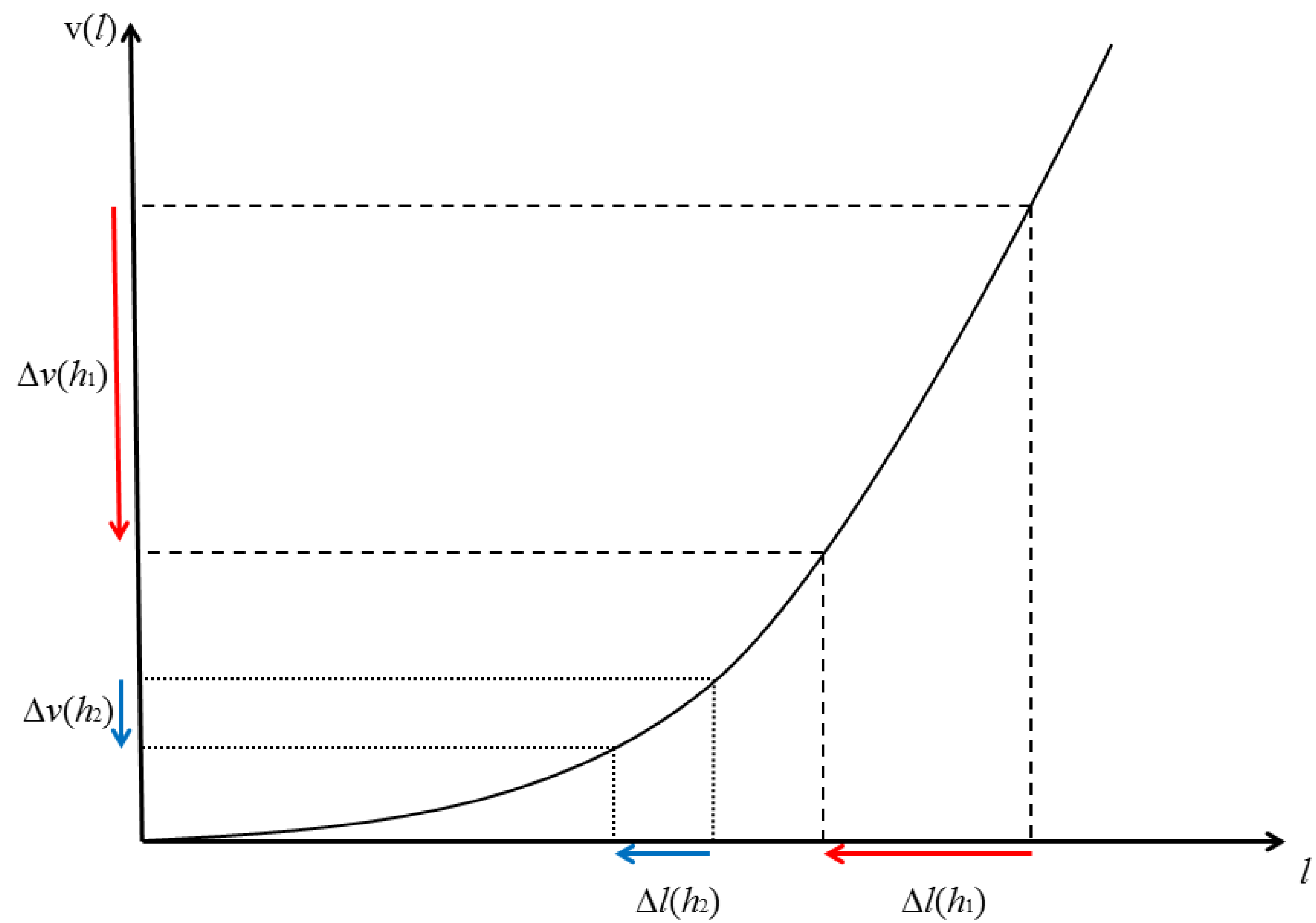

Figure 2: Decomposition of $\partial^2 v(\hat{y}, \theta, h) / (\partial\theta\partial h)$.

The first noticeable feature of Figure 2 is that an increase in h shifts the vertical lines to the left: more human capital means that a given unit of output can be produced with less labor. Since the disutility of labor is convex, it is less sensitive to changes in l for smaller l. The size of this effect is proportional to the convexity of $v(\cdot)$ and thus increasing in the elasticity parameter α.

Although human capital reduces the informational rents through its effect on labor supply, this effect may be offset because higher human capital also affects the sensitivity of labor supply with respect to changes in ability. If human capital and ability are strongly complementary for labor productivity, at higher levels of human capital h_2 much less labor is needed to produce a given output $\hat{y}$ if ability θ_2 is also higher. Visually, this means that the interval $\Delta l(h_i) \equiv l(\hat{y}, \theta_2, h_i) - l(\hat{y}, \theta_1, h_i)$ reported on the horizontal axis may be much larger at h_2 than at h_1. The figure makes clear that a sufficiently large difference between $\Delta l(h_2)$ and $\Delta l(h_1)$ can offset the labor supply effect discussed above. To sum up, when human capital is not too complementary to innate ability, $\partial^2 v(\hat{y}, \theta_2, h_2)/(\partial\theta\partial h) > 0$ so that the incentive compatibility constraint is mitigated by a marginal increase in human capital.

3 The wedges

We now compare the optimality conditions in the laissez faire and constrained-efficient allocations. In the laissez faire each family solves the maximization problem

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

where b is the bequest. Below we extend this problem by introducing a constraint that restricts assets (bequests) of families to be non-negative. For clarity we first derive results abstracting from such a constraint.

Proposition 2 *The laissez faire is characterized by the following first-order con-*

ditions 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 preferences and technologies for production and human capital accumulation such that the conditions in Proposition 2 are necessary and sufficient.[10] Then the results of Propositions 1 and 2 can be combined 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*

$$
\tau_b\left(\theta^t\right) \equiv 1-\frac{q}{\beta}\frac{\partial u(c)/\partial c}{\mathbb{E}\left[\partial u(c')/\partial c'\right]}, \tag{10}
$$

$$
\tau_l\left(\theta^t\right) \equiv 1-\frac{\partial v(y,\theta,h)/\partial y}{\partial u(c)/\partial c}, \tag{11}
$$

$$
\tau_h\left(\theta^t\right) \equiv \frac{\beta}{\frac{\partial g(h',h)}{\partial h'}}\int_{\Theta}\left[\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}}\left(\frac{\partial y'}{\partial h'}-\frac{\partial g(h'',h')}{\partial h'}\right)\right]dF(\theta')-1. \tag{12}
$$

Wedges are defined as the deviations from the laissez faire. In general, the wedges depend on the whole history of shocks since the allocation $\{c,h',y\}$ is a function of θ^t which we suppressed in the notation for convenience. In the following we denote the wedges as $\tau_j \equiv \tau_j\left(\theta^t\right)$ and the corresponding leads and lags of the wedges as $\tau_j' \equiv \tau_j'\left(\theta^{t+1}\right)$ and $\tau_{j-} \equiv \tau_{j-}\left(\theta^{t-1}\right)$, $j=b,l,h$. The wedges have a useful interpretation: constrained efficiency requires that the planner reduces (increases) bequests, labor supply or human capital, respectively, if the optimality conditions which characterize the social optimum are such that $\tau_j > 0$ ($\tau_j < 0$), $j=b,h,l$.

[10]Note that human capital is chosen for the next generation (current human capital is a state variable) so that the productivity gains from human capital accumulation accrue for the next generation and thus do not imply a direct intratemporal substitution effect for labor supply of the current generation. This timing assumption, which is plausible in our setting with families who invest into the education of their children, avoids possible non-concavities discussed in Bovenberg and Jacobs (2005), Section 2.2.

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

$$\begin{aligned} \tau_b &= 1 - \frac{1}{\mathbb{E}\left[\frac{1}{\frac{\partial u(c')}{\partial c'}}\right]\mathbb{E}\left[\frac{\partial u(c')}{\partial c'}\right]}, & (13) \\ \tau_l &= -\frac{\partial^2 v(y,\theta,h)}{\partial\theta\partial y}\frac{\mu(\theta)}{f(\theta)}, & (14) \\ \tau_h &= \Delta_l + \Delta_b + \Delta_i, & (15) \end{aligned}$$

with

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

By Jensen's inequality, we obtain the standard result that the wedge for bequests $\tau_b > 0$. The planner reduces intergenerational transfers to discourage double deviations in which parents leave bequests and their children shirk. The expression for the labor wedge τ_l is also standard: as long as ability increases productivity, $\partial^2 v(y,\theta,h)/(\partial\theta\partial y) < 0$ and it follows that $\tau_l > 0$ if the incentive constraint is binding ($\mu(\theta) > 0$). The intuition is that an additional unit of required output tightens the incentive compatibility constraint, increases the information rents and thus allows for less redistribution. Families do not internalize this effect when choosing their optimal labor supply. Corollary 2 below shows that the labor wedge in our model is analogous to the wedge in Mirrlees (1971).[11]

[11]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

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

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

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

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

Human capital wedge.—Our contribution is to derive an explicit decomposition for the human capital wedge τ_h. As shown in Proposition 3, τ_h consists of three components. The first component Δ_l relates the human capital wedge τ_h to expectations about the labor wedge τ_l'. These expectations are weighed by the marginal product of human capital. To see why τ_h positively depends on future labor wedges τ_l', note that the first-order condition for human capital in the social optimum (6) and the definition of the labor wedge (14) imply that[12]

$$\frac{\partial g(h',h)}{\partial h'} = q\int_{\Theta}\left((1-\tau_l')\frac{\partial y'}{\partial h'} - \frac{\partial g(h'',h')}{\partial h'}\right) dF(\theta') + q\int_{\Theta}\mu'(\theta')\frac{\partial^2 v(y',\theta',h')}{\partial\theta'\partial h'}d\theta'. \tag{16}$$

Since

$$\mathbb{E}\left[(1-\tau_l')\frac{\partial y'}{\partial h'}\right] = (1-\mathbb{E}[\tau_l'])\,\mathbb{E}\left[\frac{\partial y'}{\partial h'}\right] - cov\left(\tau_l', \frac{\partial y'}{\partial h'}\right),$$

equation (16) shows that the benefits from human capital accumulation for the social planner are smaller when the labor wedges τ_l' in the next generation are on average larger. If the next generation is expected to supply more labor in the laissez faire than is socially optimal ($\mathbb{E}[\tau_l'] > 0$), then the planner wants parents to invest less into the next generation's human capital which would make that generation more productive.[13] This effect is dampened if the covariance between the wedge τ_l' and the return to human capital $\partial y'/\partial h'$ is positive, thus reducing the riskiness of net returns to human capital $(1-\tau_l')\,\partial y'/\partial h'$.

[12]To derive equation (16), we have used results from the proof of Remark 1 in appendix A.2 showing that

$$\frac{\partial v(y',\theta',h')}{\partial h'} = -\frac{\partial v(y',\theta',h')}{\partial y'}\frac{\partial y'}{\partial h'}.$$

[13]In Section 4 we discuss that this result is not at odds with Bovenberg and Jacobs (2005) who show that human capital should be subsidized to offset the distortions of labor income taxation. Our findings differ from Stantcheva (2014), corollary 1, since the definition of the wedge for human capital in Stantcheva (2014) already configures a positive relationship between the wedge for human capital τ_S (notation of Stantcheva) and the wedge for labor τ_l. For the interpretation note that in Stantcheva the wedge τ_S is defined such that agents accumulate too *little* human capital in the laissez faire if $\tau_S > 0$. This case would correspond to $\tau_h < 0$ in our model.

The second component Δ_b relates the wedge for human capital to the wedge for bequests τ_b. The first term in Δ_b

$$\frac{1}{\frac{\partial g(h',h)}{\partial h'}}\mathbb{E}\left[\beta\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}}-q\right]\mathbb{E}\left[\frac{\partial y'}{\partial h'}-\frac{\partial g(h'',h')}{\partial h'}\right]$$

is positive:

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

and the marginal product of human capital

$$\frac{\partial y'}{\partial h'}-\frac{\partial g(h'',h')}{\partial h'}>0$$

because $\partial g(h'',h')/\partial h' < 0$, as shown in Remark 1 in appendix A.2. Intuitively, if the planner discourages bequests, constrained efficiency also involves discouraging human capital accumulation since education 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 Δ_b 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, $\partial y'/\partial h' - \partial g(h'',h')/\partial h'$, and consumption of the next generation increase with ability θ'. In this case, the planner is less inclined to discourage human-capital investment because its returns are riskier.

The components Δ_l and Δ_b show that the wedge for human capital accumulation depends on the wedges for labor supply and bequests which is intuitive since education alters the marginal product of labor and transfers resources across periods. The last component Δ_i captures the effect of human capital accumulation on the incentive-compatibility constraint. Δ_i depends on the covariance between the impact 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.

The incentive wedge Δ_i for human capital.—In order to understand the economic intuition captured by Δ_i, it is useful to rewrite it as[14]

$$\Delta_i=-\frac{q}{\frac{\partial g(h',h)}{\partial h'}}\int_{\Theta}\mu'\left(\theta'\right)\frac{\partial^2 v(y',\theta',h')}{\partial\theta'\partial h'}d\theta' \,. \tag{17}$$

[14] See the proof of Proposition 3 for a derivation of equation (17).

As explained in Section 2.3, the integral on the right hand side measures by how much a marginal increase in human capital is expected to relax the incentive-compability constraint for the next generation. This benefit is ignored by the families because they take the allocation as given and thus ignore the impact that their human-capital investments have on the incentive-compatibility constraint. Families do not internalize the effect of human capital on the informational rents, which drives a wedge between the optimal choice of the family and that of the planner. Whether this wedge is positive or negative depends on the elasticity of labor supply and complementarity between human capital and ability.

Corollary 3 *Under assumptions **[A1']** and **[A2]**, $\partial^2 v(y',\theta',h')/(\partial\theta'\partial h') > 0$ if and only if $\chi > -\alpha$. Then $\Delta_i < 0$, showing that the planner has a motive to increase human capital accumulation in order to relax the incentive compatibility constraint.*

Informational rents provide a rationale for educational subsidies when $\chi < -\alpha$. For a plausible Frisch elasticity of labor supply of $\varepsilon = 0.5$, $\alpha = \varepsilon^{-1} + 1 = 3$ (Chetty, 2012, Table III). In order to gauge whether a value of $\chi > -3$ is empirically plausible, we refer to empirical results by Cunha et al. (2006) who report in Table 3, p. 748, that the marginal return to college, in percentage terms, of an individual from 5th percentile of the ability distribution (according to the Armed Forces Qualification Test) is 50% of the return of an individual from the 95th percentile. In competitive labor markets this marginal return corresponds to increases in productivity. In terms of our model

$$\frac{\frac{\frac{\partial A(\theta,h)}{\partial h}}{A(\theta,h)}|_{\theta=\theta_5}}{\frac{\frac{\partial A(\theta,h)}{\partial h}}{A(\theta,h)}|_{\theta=\theta_{95}}} = 0.5,$$

where θ_j denotes the j-th percentile of the distribution of ability θ. Given assumption **[A2]**, there is a simple relationship between the ratio of percentage increases in productivity and the ratio of productivity levels which depends on the complementarity parameter χ:

$$\frac{\frac{\frac{\partial A(\theta,h)}{\partial h}}{A(\theta,h)}|_{\theta=\theta_5}}{\frac{\frac{\partial A(\theta,h)}{\partial h}}{A(\theta,h)}|_{\theta=\theta_{95}}} = \left(\frac{A(\theta,h)|_{\theta=\theta_5}}{A(\theta,h)|_{\theta=\theta_{95}}}\right)^{-\chi}. \tag{18}$$

In the Cobb-Douglas case, $\chi = 0$ so that the ratio of productivity increases is one. With $A(\theta,h)|_{\theta=\theta_5} < A(\theta,h)|_{\theta=\theta_{95}}$ and

$$\frac{\frac{\frac{\partial A(\theta,h)}{\partial h}}{A(\theta,h)}|_{\theta=\theta_5}}{\frac{\frac{\partial A(\theta,h)}{\partial h}}{A(\theta,h)}|_{\theta=\theta_{95}}} < 1,$$

as reported by Cunha et al. (2006), $\chi < 0$: ability θ and human capital h are more complementary than in the Cobb-Douglas case.

Equation (18) and the estimates by Cunha et al. (2006) imply that $\chi \geq -3$ if the ratio of productivity levels

$$\frac{A(\theta, h)\,|_{\theta=\theta_5}}{A(\theta, h)\,|_{\theta=\theta_{95}}} \leq 0.5^{\frac{1}{3}} \approx 0.79.$$

Hanushek (2006), p. 871, reports that an increase of standardized AFQT test scores by one standard deviation increases earnings by 12%. Using AFQT test scores to approximate ability and assuming that ability is normally distributed, this implies $A(\theta, h)\,|_{\theta=\theta_5}/A(\theta, h)\,|_{\theta=\theta_{95}} \approx 2/3$ for individuals working the same amount of hours with average productivity A normalized to one. Then

$$\chi = -\frac{\ln(0.5)}{\ln(2/3)} \approx -1.71.$$

This suggests that $\chi \geq -3$ is empirically plausible so that human capital accumulation relaxes the incentive compatibility constraint.

Comparison with the literature.— Our finding that human capital alleviates incentive constraints for empirically plausible parameters differs from Stantcheva (2014) since we consider the effect of human capital accumulation on incentive compatibility through changes in labor supply. This effect is important because the planner can choose output but cannot directly control the labor supply of families. Thus, optimal redistribution has to take into account how human-capital investment affects incentives to produce a given level of output by changing labor supply. Considering this effect we find that human capital mitigates the incentive problem if $\chi > -\alpha$, instead of $\chi > 0$ as in Stantcheva (2014).

3.1 Liquidity constraints

In this subsection we show how our results modify if we impose the constraint that assets (bequests) of a family cannot be negative. In this case parents cannot require children to make transfers to them and children within a family cannot take on debt obligations.

In the laissez faire each family solves the maximization problem

$$
\begin{aligned}
W(\theta,b,h,t) &= \max_{\{b',h',l\}}\left\{\mathbf{U}(c,l)+\beta\int_{\Theta}W(\theta',b',h',t+1)\,dF(\theta')\right\}\\
\text{s.t. } b' &= (1+r)b-c-e+y,\\
b' &\geq 0,\\
y &= Y(h,\theta,l),\\
h' &= h'(e,h)\text{ so that } e=g(h',h),
\end{aligned}
$$

where $b'\geq 0$ imposes that bequests cannot be negative and the multiplier $\eta>0$ if this constraint is binding.

Proposition 4 *If bequests are required to be non-negative, the laissez faire is characterized by the following first-order conditions for bequests, human capital and labor supply:*

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

The modified definitions of the wedges are as follows:

Definition 2 *If bequests are required to be non-negative, the wedges for bequests τ_b^c, labor supply τ_l^c and human capital τ_h^c are*

$$
\tau_b^c \equiv 1-\frac{q}{\beta}\frac{\partial u(c)/\partial c-\eta}{\mathbb{E}\left[\partial u(c')/\partial c'\right]}, \tag{19}
$$

$$
\tau_l^c \equiv 1-\frac{\partial v(y,\theta,h)/\partial y}{\partial u(c)/\partial c-\eta}, \tag{20}
$$

$$
\tau_h^c \equiv \frac{\beta}{\frac{\partial g(h',h)}{\partial h'}}\int_{\Theta}\left[\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}-\eta}\left(\frac{\partial y'}{\partial h'}-\frac{\partial g(h'',h')}{\partial h'}\right)\right]dF(\theta')-1. \tag{21}
$$

Combining the results of Propositions 1 and 4, we then find:

Proposition 5 *If bequests are required to be non-negative, the first-order conditions of the planner's problem imply under assumption **[A1]** that*

$$\tau_b^c = 1 - \frac{1}{\mathbb{E}\left[\frac{1}{\frac{\partial u(c')}{\partial c'}}\right]\mathbb{E}\left[\frac{\partial u(c')}{\partial c'}\right]} + \frac{\eta}{\frac{\beta}{q}\mathbb{E}\left[\frac{\partial u(c')}{\partial c'}\right]}, \tag{22}$$

$$\tau_l^c = -\frac{\mu(\theta)}{f(\theta)}\frac{\partial^2 v(y,\theta,h)}{\partial\theta\partial y} - \frac{\eta}{\frac{\partial u(c)}{\partial c} - \eta}\frac{\frac{\partial v(y,\theta,h)}{\partial y}}{\frac{\partial u(c)}{\partial c}}, \tag{23}$$

$$\tau_h^c = \Delta_l^c + \Delta_b^c + \Delta_i^c + \Delta_c, \tag{24}$$

with

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

Thus, if the liquidity constraint for a family is binding ($\eta > 0$), the wedge on labor decreases ceteris paribus as the planner encourages more labor earnings to alleviate the constraint. The wedge for savings and human capital increase ceteris paribus since a binding liquidity constraint implies that the future generation has more resources than would be socially optimal.

4 Implementation of the constrained-efficient allocation

An important question is how the socially optimal allocation of the planner's problem can be implemented in a decentralized economy. Stantcheva (2014), Section 6, shows how to achieve this with different combinations of instruments. One possibility is to rely on loans for human capital accumulation with payments that are contingent on the history of loans and income. The socially optimal allocation is then implemented if these history-dependent loan repayments are combined with taxes on labor income and bequests that condition only on current income and current bequests, respectively.

Stantcheva (2014) also applies results of Albanesi and Sleet (2006) to show that, with i.i.d. ability θ, the history dependence of the tax system becomes much simpler since the history can be summarized by two state variables: bequests and human capital. It is then possible to implement the constrained efficient allocation, for example, either with means-tested grants that depend on labor income y, human capital accumulation h' and condition on the initial state variables b and h; or with loans for human capital accumulation featuring repayment schedules that depend on y, h', condition on b and h and are complemented with labor income taxes which only depend on current income y.

For concreteness, let us illustrate how marginal taxes for labor and human capital relate to the respective wedges (see Stantcheva, 2014, and Kocherlakota, 2010, ch. 5, for how taxes on bequests implement the savings wedge). Assume the tax schedule $T(b,h,y,h')$ so that agents solve the maximization problem

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

The first-order condition for labor supply and the definition of the labor wedge (11) imply that the labor wedge equals the marginal income tax: $\tau_l = \partial T(\cdot)/\partial y$. The first-order condition for human capital

$$\begin{aligned} \left(\frac{\partial g(h',h)}{\partial h'}+\frac{\partial T(\cdot)}{\partial h'}\right)\frac{\partial \mathbf{U}(c,l)}{\partial c} &= \beta\int_{\Theta}\left[\frac{\partial y'}{\partial h'}\left(1-\frac{\partial T'(\cdot)}{\partial y'}\right)\frac{\partial \mathbf{U}(c',l')}{\partial c'}\right]dF(\theta') \\ &\quad -\beta\int_{\Theta}\left[\left(\frac{\partial g(h'',h')}{\partial h'}+\frac{\partial T'(\cdot)}{\partial h'}\right)\frac{\partial \mathbf{U}(c',l')}{\partial c'}\right]dF(\theta') \end{aligned}$$

and the definition of the wedge for human capital (15) imply

$$\frac{\partial T(\cdot)}{\partial h'}=\frac{\partial g(h',h)}{\partial h'}\tau_h-\beta\int_{\Theta}\left[\left(\frac{\partial y'}{\partial h'}\frac{\partial T'(\cdot)}{\partial y'}+\frac{\partial T'(\cdot)}{\partial h'}\right)\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}}\right]dF(\theta'). \qquad (25)$$

As pointed out by Stantcheva (2014), a positive wedge for human capital does not necessarily imply a positive current marginal tax on human capital accumulation in a dynamic model. The second term on the right-hand side shows that this also depends on how the marginal tax rates in the next period are correlated with the marginal utility of consumption. If the change of the marginal tax rate resulting from additional human capital accumulation, $\frac{\partial y'}{\partial h'}\frac{\partial T'(\cdot)}{\partial y'}+\frac{\partial T'(\cdot)}{\partial h'}$, is negatively correlated with marginal utility of consumption, human capital hedges consumption risk. Then the current marginal tax on human capital accumulation $\partial T(\cdot)/\partial h'$ can be positive although τ_h is negative.

Equation (25) allows us to relate our results further to Bovenberg and Jacobs (2005) who show that human capital should be subsidized if taxation of labor income distorts the decision to accumulate human capital. We recover the analogon of this result in our model, if human capital accumulation is socially optimal in the laissez faire without tax distortions as in Bovenberg and Jacobs (2005). Then $\tau_h = 0$ in equation (25) and locally

$$\begin{aligned}\frac{\partial T(\cdot)}{\partial h'} &= -\beta\int_{\Theta}\left[\left(\frac{\partial y'}{\partial h'}\frac{\partial T'(\cdot)}{\partial y'}+\frac{\partial T'(\cdot)}{\partial h'}\right)\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}}\right]dF(\theta') \\ &= -\beta\mathbb{E}\left[\frac{\partial y'}{\partial h'}\frac{\partial T'(\cdot)}{\partial y'}+\frac{\partial T'(\cdot)}{\partial h'}\right]\mathbb{E}\left[\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}}\right] \\ &\quad -\beta cov\left(\frac{\partial y'}{\partial h'}\frac{\partial T'(\cdot)}{\partial y'}+\frac{\partial T'(\cdot)}{\partial h'},\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}}\right).\end{aligned}$$

Thus, the current marginal tax on human capital accumulation $\partial T(\cdot)/\partial h'$ is negatively related to the expected tax change for the next generation, $\mathbb{E}\left[\frac{\partial y'}{\partial h'}\frac{\partial T'(\cdot)}{\partial y'}+\frac{\partial T'(\cdot)}{\partial h'}\right]$, resulting from an additional marginal unit of human capital. Compared with Bovenberg and Jacobs (2005), this tax change does not only consist of the additional marginal income tax but also of the change of taxes due to the higher human capital stock of the next generation. Moreover, the returns to human capital are uncertain in our model so that it matters whether the tax changes reduce consumption risk. Human capital accumulation should then be subsidized if $\mathbb{E}\left[\frac{\partial y'}{\partial h'}\frac{\partial T'(\cdot)}{\partial y'}+\frac{\partial T'(\cdot)}{\partial h'}\right] > 0$ and the future tax changes caused by human capital accumulation do not reduce consumption risk too much (i.e., the covariance is not too negative).

This is not the whole story in our model, however, since $\tau_h \neq 0$ in the laissez faire without tax distortions because of the incentive effect of human capital accumulation derived in Proposition 3. Thus, we have an additional term in equation (25). The incentive effect is not present in Bovenberg and Jacobs (2005) since the assumption of a Cobb-Douglas production function implies $\chi = 0$ and the effect of human capital on labor supply in the incentive compatibility constraint is not considered. Making explicit that the planner chooses output but cannot choose labor supply directly reveals this effect in our model.

Existing tax and subsidy systems for student loans in continental Europe and Anglo-Saxon countries contain elements which resemble the tax schedules analyzed above. These schedules condition on bequests and human capital which roughly corresponds to grants or repayment schedules for student loans in the real world that condition on parents' permanent income (which is highly correlated with human capital) and parents' wealth (which is correlated with bequests). We have discussed above that the relatively simple tax schedules $T(b, h, y, h')$, which condition on all relevant history through the state variables b and h, allow to implement the social optimum if ability θ is i.i.d. An interesting question for further research is how large the persistence of ability across generations has to be so that the simple tax and subsidy schedules observed in reality imply sizable deviations from the social optimum and thus substantial welfare losses.

5 Conclusion

We have shown that human capital investment by families is not constrained efficient if the ability of generations in a family dynasty 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, the wedge for bequests and an incentive term. 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. Since the planner's Hamiltonian reads

$$\begin{aligned}\mathcal{H} &= [c(\omega(\theta)-\beta V'(\theta), y(\theta), \theta, h) + g(h'(\theta), h) - y(\theta) + q\Gamma(V'(\theta), h'(\theta), t+1)] f(\theta) \\ &\quad +\lambda [V - \omega(\theta) f(\theta)] \\ &\quad +\mu(\theta) [\partial U (c(\omega(\theta)-\beta V'(\theta), y(\theta), \theta, h), y(\theta), \theta, h)/\partial\theta],\end{aligned}$$

the first-order conditions are

$$\left[\frac{\partial c(\theta)}{\partial V'(\theta)} + q\frac{\partial \Gamma(V'(\theta), h'(\theta), t+1)}{\partial V'(\theta)}\right] f(\theta) = -\mu(\theta)\frac{\partial^2 U(\cdot)}{\partial\theta\partial c(\theta)}\frac{\partial c(\theta)}{\partial V'(\theta)}, \quad (26)$$

$$\left[\frac{\partial g(h'(\theta), h)}{\partial h'(\theta)} + q\frac{\partial \Gamma(V'(\theta), h'(\theta), t+1)}{\partial h'(\theta)}\right] f(\theta) = 0, \quad (27)$$

$$\mu(\theta)\left[\frac{\partial^2 U(\cdot)}{\partial\theta\partial c(\theta)}\frac{\partial c(\theta)}{\partial y(\theta)} + \frac{\partial^2 U(\cdot)}{\partial\theta\partial l(\theta)}\frac{\partial l(\theta)}{\partial y(\theta)}\right] = -\left[\frac{\partial c(\theta)}{\partial y(\theta)} - 1\right] f(\theta). \quad (28)$$

The costate variable satisfies

$$\frac{\partial\mu(\theta)}{\partial\theta} = -\left[\frac{\partial c(\theta)}{\partial\omega(\theta)} - \lambda + \frac{\mu(\theta)}{f(\theta)}\frac{\partial^2 U(\cdot)}{\partial\theta\partial c(\theta)}\frac{\partial c(\theta)}{\partial\omega(\theta)}\right] f(\theta); \quad (29)$$

with the usual boundary conditions $\lim_{\theta\to\underline{\theta}}\mu(\theta) = 0$ and $\lim_{\theta\to\bar{\theta}}\mu(\theta) = 0$. We use assumption **[A1]** to invert the utility function

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

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 (26) simplifies to

$$\frac{1}{\partial u(c(\theta))/\partial c(\theta)} = \frac{q}{\beta}\frac{\partial\Gamma(V'(\theta), h'(e(\theta), h), t)}{\partial V'(\theta)} = \frac{q}{\beta}\lambda'(\theta),$$

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

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 (28) 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 θ and that $\partial^2 U(\cdot)/(\partial c(\theta)\partial\theta) = 0$:

$$\begin{aligned}\frac{\partial\Gamma(V,h,t)}{\partial h} &= \int_\Theta \left(\frac{\partial c(\theta)}{\partial h} + \frac{\partial g(h'(\theta),h)}{\partial h}\right) dF(\theta) + \int_\Theta \mu(\theta)\frac{\partial^2 U(\cdot)}{\partial\theta\partial h}d\theta \\ &= \int_\Theta \left(\frac{\partial v(y(\theta),\theta,h)/\partial h}{\partial u(c(\theta))/\partial c(\theta)} + \frac{\partial g(h'(\theta),h)}{\partial h}\right) dF(\theta) - \int_\Theta \mu(\theta)\frac{\partial^2 v(y(\theta),\theta,h)}{\partial\theta\partial h}d\theta.\end{aligned}$$

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 (27) and (28). The envelope condition for human capital can then be inserted into the optimality condition for human capital (27) to obtain

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

For $\partial^2 U(\cdot)/(\partial c(\theta)\partial\theta) = 0$, equation (29) implies

$$\mu(\theta) = \int_{\underline{\theta}}^{\theta}\left[-\frac{1}{\partial u(c(x))/\partial c(x)} + \lambda\right] dF(x). \tag{30}$$

■

Remark 1 *Assume that the technology for human capital accumulation is given by*

[A3]: $h'(e,h) = [\phi e^\rho + (1-\phi)h^\rho]^{1/\rho}$ ***with*** $\rho \in (-\infty, 1]$ ***and*** $\phi \in (0,1)$.

*Under assumptions **[A1']**, **[A2]** and **[A3]**, we have*

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

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

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

Differentiating these expressions a second time, we get

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

and

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

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

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

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

$$e = g(h', h) = \max\left\{\left[\frac{1}{\phi}(h')^{\rho} - \frac{1-\phi}{\phi}h^{\rho}\right]^{1/\rho}; 0\right\}.$$

The expenditure function $g(h', h)$ has constant returns to scale in h' and h at an interior solution where

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

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

■

Corollary 1
Follows immediately from Remark 1.

Proposition 2
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 intratemporal 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

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

■

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 (7) of the planner. To derive the analogous expression for τ_b, we recall that $\lambda'(\theta)=\mathbb{E}\left[\frac{\partial c'(\theta')}{\partial u(c'(\theta'))}\right]$ and rearrange the definition of τ_b to substitute $\partial u(c)/\partial c$ in condition (5). 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 (6):

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

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

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

which can be rearranged to

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

The first term equals Δ_l using the definition of the labor wedge (11). The second term equals Δ_b using that $\mathbb{E}\left(xy\right)=cov(x,y)+\mathbb{E}\left(x\right)\mathbb{E}\left(y\right)$.

In the remaining part of the proof, we focus on the last term of τ_h to derive Δ_i. Integrating the integral of the last term by parts,

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

The first term on the right-hand side is equal to zero because of the boundary conditions for $\mu'\left(\theta'\right)$. Thus, using (29) imposing assumption **[A1]**, the last term of the wedge τ_h becomes

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

Since by (5),

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

we get

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

The integral simplifies since it is equivalent to

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

where the second equality follows from the reciprocal Euler equation

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

This concludes the proof. ∎

Corollary 2
Proof. To compare the labor wedge in our model with the literature, we use definition (11) to derive

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

Thus, (14) implies that at the solution of the planner's problem,

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

By Remark 1,

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

where we have substituted in $\mu(\theta)$ using (30). ■

Corollary 3
Follows immediately from Remark 1 and Proposition 3.

Proposition 4
Proof. *Savings.* The first-order condition for savings reads

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

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

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

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

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

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

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

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

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

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

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

■

Proposition 5

Proof. We derive the wedge τ_l^c evaluated at the solution of the planner's problem using the definition for τ_l^c in the first-order condition (7) of the planner. Condition (7) implies

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

which, using the definition of the wedge τ_l^c, becomes

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

Simplifying, we get

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

where $\frac{\partial u(c)}{\partial c} - \eta > 0$ since $\int_{\Theta} \frac{\partial W(\theta',b',h')}{\partial b'} dF(\theta') > 0$. To derive the analogous expression for τ_b, we recall that $\lambda'(\theta) = \mathbb{E}\left[\frac{1}{\frac{\partial u(c')}{\partial c'}}\right]$ and rearrange the definition of τ_b^c to substitute $\partial u(c)/\partial c$ in condition (5). Condition (5) implies

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

The definition of the wedge τ_b^c can be rearranged to

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

so that substituting out $\partial u\left(c(\theta)\right)/\partial c(\theta)$ yields

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

Solving this expression for τ_b^c results in

$$\tau_b^c = 1 - \frac{1}{\mathbb{E}\left[\frac{1}{\frac{\partial u(c'))}{\partial c'}}\right]\mathbb{E}\left[\partial u\left(c'\right))/\partial c'\right]} + \frac{\eta}{\frac{\beta}{q}\mathbb{E}\left[\partial u\left(c'\right)/\partial c'\right]}$$

The wedge for human capital implied by the solution to the planner's problem is obtained by adding τ_h^c on both sides of condition (6):

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

Substituting in the definition of the wedge τ_h^c on the right-hand side, we get

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

which, using

$$\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}-\eta} = \frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}} + \frac{\eta}{\frac{\partial u(c)}{\partial c}-\eta}\frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}},$$

can be rearranged to

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

The first term equals Δ_l^c using the definition of the labor wedge (11). The second term equals Δ_b^c using that $\mathbb{E}(xy) = cov(x,y) + \mathbb{E}(x)\mathbb{E}(y)$. The third term of τ_h^c can be shown to yield Δ_i^c, as derived in the proof of Proposition 3. The fourth term equals Δ_c. ■

A.3 Persistent types

We now turn our attention to the general case where types are correlated from one generation to the next. The analysis with persistent types draws on results by Kapička (2013), applied to dynamic optimal taxation problems by Farhi and Werning (2013), Golosov et al. (2013) and Stantcheva (2014). Following Pavan, Segal and Toikka (forthcoming), the envelope condition in the problem with persistent shocks is:

$$
\frac{\partial \omega(\theta)}{\partial \theta} = \frac{\partial U(c,y,\theta,h)}{\partial \theta} + \beta \int_\Theta \omega(\theta') \frac{\partial f(\theta'|\theta)}{\partial \theta} d\theta'. \tag{31}
$$

This condition serves as local incentive compatibility constraint in the relaxed problem based on the first-order approach. The recursive formulation with persistent types requires that Δ and V are treated as state variables where

$$
\Delta(\theta) \equiv \int_\Theta \omega(\theta) \frac{\partial f(\theta|\theta_-)}{\partial \theta_-} d\theta,
$$

so that

$$
\frac{\partial \omega(\theta)}{\partial \theta} = \frac{\partial U(c,y,\theta,h)}{\partial \theta} + \beta \Delta'.
$$

As before we consider the relaxed planner's problem, with local constraints evaluated at the truthful equilibrium reports, and apply optimal control techniques. The recursive problem is

$$\begin{aligned}
& \Gamma\left(V, \Delta, \theta_{-}, h, t\right) \qquad (32)\\
= & \min_{\left\{c, y, h', \Delta', V'\right\}}\left\{\int_{\Theta}\left[c+g(h', h)-y(\theta)+q \Gamma\left(V', \Delta', \theta, h', t+1\right)\right] dF\left(\theta \mid \theta_{-}\right)\right\} \\
s.t.\, \omega(\theta) = & U(c, y, \theta, h)+\beta V', \\
V = & \int_{\Theta} \omega(\theta)\, dF(\theta), \\
\Delta = & \int_{\Theta} \omega(\theta) \frac{\partial f\left(\theta \mid \theta_{-}\right)}{\partial \theta_{-}} d\theta, \\
\frac{\partial \omega(\theta)}{\partial \theta} = & \frac{\partial U(c, y, \theta, h)}{\partial \theta}+\beta \Delta'.
\end{aligned}$$

As before, we substitute consumption with the promise-keeping constraint, defining consumption $c\left(\omega(\theta)-\beta V', y, \theta, h\right)$ as an implicit function of other control and state variables. This enables us to write the Hamiltonian associated with the planner's problem as

$$\begin{aligned}
\mathcal{H} = & \left[c\left(\omega(\theta)-\beta V', y, \theta, h\right)+g(h', h)-y+q \Gamma\left(V', \Delta', \theta, h', t+1\right)\right] f\left(\theta \mid \theta_{-}\right) \\
& +\lambda(\theta_{-})\left[V-\omega(\theta) f\left(\theta \mid \theta_{-}\right)\right]+\gamma(\theta_{-})\left[\Delta-\omega(\theta) \frac{\partial f\left(\theta \mid \theta_{-}\right)}{\partial \theta_{-}}\right] \\
& +\mu(\theta)\left[\frac{\partial U\left(c\left(\omega(\theta)-\beta V, y, \theta, h\right), y, \theta, h\right)}{\partial \theta}+\beta \Delta'\right] .
\end{aligned}$$

The costate variable satisfies

$$\frac{\partial \mu(\theta)}{\partial \theta}=-\left[\frac{1}{\partial u(c) / \partial c}-\lambda(\theta_{-})-\gamma(\theta_{-}) \frac{\frac{\partial f\left(\theta \mid \theta_{-}\right)}{\partial \theta_{-}}}{f\left(\theta \mid \theta_{-}\right)}+\frac{\mu(\theta)}{f\left(\theta \mid \theta_{-}\right)} \frac{\partial^{2} U(\cdot)}{\partial \theta \partial c} \frac{\partial c}{\partial \omega(\theta)}\right] f\left(\theta \mid \theta_{-}\right), \qquad (33)$$

with $\lim_{\theta \rightarrow \underline{\theta}} \mu(\theta)=0$ and $\lim_{\theta \rightarrow \bar{\theta}} \mu(\theta)=0$. The first-order conditions read

$$\begin{aligned}
\frac{\partial \mathcal{H}(\cdot)}{\partial V'} = & \left[\frac{\partial c}{\partial V'}+q \frac{\partial \Gamma\left(V', \Delta', \theta, h', t+1\right)}{\partial V'}\right] f\left(\theta \mid \theta_{-}\right)+\mu(\theta) \frac{\partial^{2} U(\cdot)}{\partial \theta \partial c} \frac{\partial c}{\partial V'}=0, \\
\frac{\partial \mathcal{H}(\cdot)}{\partial \Delta'} = & \left[q \frac{\partial \Gamma\left(V', \Delta', \theta, h', t+1\right)}{\partial \Delta'}\right] f\left(\theta \mid \theta_{-}\right)+\beta \mu(\theta)=0, \\
\frac{\partial \mathcal{H}(\cdot)}{\partial y} = & \left[\frac{\partial c}{\partial y}-1\right] f\left(\theta \mid \theta_{-}\right)+\mu(\theta)\left[\frac{\partial^{2} U(\cdot)}{\partial \theta \partial c} \frac{\partial c}{\partial y}+\frac{\partial^{2} U(\cdot)}{\partial \theta \partial l} \frac{\partial l}{\partial y}\right]=0, \\
\frac{\partial \mathcal{H}(\cdot)}{\partial h'} = & \frac{\partial g(h', h)}{\partial h'}+q \frac{\partial \Gamma\left(V', h', t+1\right)}{\partial h'}=0 .
\end{aligned}$$

For the optimality condition for human capital, we use the envelope condition

$$\begin{aligned}\frac{\partial\Gamma\left(V,\Delta,\theta_{-},h,t\right)}{\partial h} &= \int_{\Theta}\left(\frac{\partial c}{\partial h}+\frac{\partial g(h',h)}{\partial h}\right)dF(\theta|\,\theta_{-})+\int_{\Theta}\mu(\theta)\frac{\partial^{2}U\left(\cdot\right)}{\partial\theta\partial h}d\theta \\ &= \int_{\Theta}\left(\frac{\partial v(y,\theta,h)/\partial h}{\partial u\left(c\right)/\partial c}+\frac{\partial g(h',h)}{\partial h}\right)dF(\theta|\,\theta_{-})-\int_{\Theta}\mu(\theta)\frac{\partial^{2}v(y,\theta,h)}{\partial\theta\partial h}d\theta \\ &\quad +\int_{\Theta}\mu(\theta)\frac{\partial^{2}u\left(\cdot\right)}{\partial\theta\partial c}\frac{\partial c}{\partial h}d\theta\,.\end{aligned}$$

Imposing **[A1]** and using the envelope conditions $\partial\Gamma\left(\cdot\right)/\partial V=\lambda(\theta_{-})$ and $\partial\Gamma\left(\cdot\right)/\partial\Delta=\gamma(\theta_{-})$ allows us to derive the system of first-order conditions analogous to Proposition 1 but with persistent types:

$$\frac{\partial\mathcal{H}\left(\cdot\right)}{\partial V'} = \left[-\frac{\beta}{\partial u\left(c\left(\theta\right)\right)/\partial c(\theta)}+q\lambda'(\theta)\right]f(\theta|\,\theta_{-})=0\,, \tag{34}$$

$$\frac{\partial\mathcal{H}\left(\cdot\right)}{\partial\Delta'} = q\gamma'\left(\theta\right)f(\theta|\,\theta_{-})+\beta\mu\left(\theta\right)=0\,, \tag{35}$$

$$\frac{\partial\mathcal{H}\left(\cdot\right)}{\partial y} = \left[\frac{\partial v(y,\theta,h)/\partial y}{\partial u\left(c\right)/\partial c}-1\right]f(\theta|\,\theta_{-})-\frac{\partial^{2}v(y,\theta,h)}{\partial\theta\partial y}\mu\left(\theta\right)=0\,, \tag{36}$$

$$\begin{aligned}\frac{\partial\mathcal{H}\left(\cdot\right)}{\partial h'} &= \frac{\partial g(h',h)}{\partial h'}+q\int_{\Theta}\left(\frac{\frac{\partial v(y',\theta',h')}{\partial h'}}{\frac{\partial u(c')}{\partial c'}}+\frac{\partial g(h'',h')}{\partial h'}\right)dF(\theta'|\,\theta) \\ &- q\int_{\Theta}\mu'\left(\theta'\right)\frac{\partial^{2}v(y',\theta',h')}{\partial\theta'\partial h'}d\theta'=0.\end{aligned} \tag{37}$$

The system of equations is similar to the system derived for i.i.d. types but note that persistence of types alters the multiplier of the incentive compatibility constraint $\mu\left(\theta\right)$. Using (33) to substitute out $\mu(\theta)$ in equation (36), we get

$$\left[\frac{\frac{\partial v(y,\theta,h)}{\partial y}}{\frac{\partial u(c)}{\partial c}}-1\right]f(\theta|\,\theta_{-})=\frac{\partial^{2}v(y,\theta,h)}{\partial\theta\partial y}\int_{\underline{\theta}}^{\theta}\left[-\frac{1}{\frac{\partial u(c(x))}{\partial c(x)}}+\lambda(\theta_{-})+\gamma(\theta_{-})\frac{\frac{\partial f(\,\theta|\theta_{-})}{\partial\theta_{-}}}{f(x|\,\theta_{-})}\right]f(x|\,\theta_{-})dx.$$

As in Farhi and Werning (2013), we can then characterize the dynamics of the labor wedge.

Proposition 6 *Under assumptions [A1], [A1'] and [A2], the labor wedge satisfies* *the following condition*

$$\begin{aligned}&\int_{\Theta}\left(\frac{\tau_l^p}{1-\tau_l^p}\right)\frac{q}{\beta}\frac{\partial u\left(c_{-}\right)/\partial c_{-}}{\partial u\left(c\right)/\partial c}\pi\left(\theta\right)f(\theta|\,\theta_{-})d\theta \\ = & \frac{\tau_{l-}^p}{1-\tau_{l-}^p}\theta_{-}^{1-\chi}\int_{\Theta}\Pi\left(\theta\right)\frac{\partial f(\theta|\,\theta_{-})}{\partial\theta_{-}}d\theta \\ &+\frac{\alpha\xi}{A^{\chi}}\int_{\Theta}\Pi\left(\theta\right)\left[\frac{q}{\beta}\frac{\partial u\left(c_{-}\right)/\partial c_{-})}{\partial u(c)/\partial c}-1\right]f(\theta|\,\theta_{-})d\theta\end{aligned}$$

for any given function $\pi\left(\theta\right)$ *with* $\Pi\left(\theta\right)$ *being the primitive of* $\pi\left(\theta\right)/\theta^{1-\chi}$*.*

Proof. Under assumptions [**A1**], **[A1']** and **[A2]**, the derivatives in the proof of Remark 1 imply

$$\frac{\partial^2 v(y,\theta,h)}{\partial\theta\partial y}=-\alpha\frac{\xi\theta^{\chi-1}}{A^{\chi}}\frac{\partial v\left(l\right)}{\partial l}\frac{1}{A}=-\alpha\frac{\xi\theta^{\chi-1}}{A^{\chi}}\frac{\partial v(y,\theta,h)}{\partial y}\,.$$

Substituting this equality in (36) and using the definition of the labor wedge as before, we obtain

$$1-\frac{\frac{\partial v(y,\theta,h)}{\partial y}}{\frac{\partial u(c)}{\partial c}}=-\frac{\mu\left(\theta\right)}{f(\theta|\,\theta_{-})}\frac{\partial^2 v(y,\theta,h)}{\partial\theta\partial y}=\frac{\mu\left(\theta\right)}{f(\theta|\,\theta_{-})}\alpha\frac{\xi\theta^{\chi-1}}{A^{\chi}}\frac{\partial v(y,\theta,h)}{\partial y}\,.$$

or, using the definition of the labor wedge (11),

$$\left[\frac{\frac{\partial u(c)}{\partial c}}{\frac{\partial v(y,\theta,h)}{\partial y}}-1\right]\frac{f(\theta|\,\theta_{-})}{\frac{\partial u(c(\theta))}{\partial c(\theta)}}=\left[\frac{\tau_l^p}{1-\tau_l^p}\right]\frac{f(\theta|\,\theta_{-})}{\frac{\partial u(c(\theta))}{\partial c(\theta)}}=\mu\left(\theta\right)\alpha\frac{\xi\theta^{\chi-1}}{A^{\chi}}\,. \tag{38}$$

This identity implies that for any differentiable function $\pi\left(\theta\right)$

$$\int_{\Theta}\left(\frac{\tau_l^p}{1-\tau_l^p}\right)\frac{\pi\left(\theta\right)}{\frac{\partial u(c)}{\partial c}}f(\theta|\,\theta_{-})d\theta=\frac{\alpha\xi}{A^{\chi}}\int_{\Theta}\mu\left(\theta\right)\frac{\pi\left(\theta\right)}{\theta^{1-\chi}}d\theta\,. \tag{39}$$

Integrating by part the expression on the right-hand side and denoting with $\Pi\left(\theta\right)$ the primitive of $\pi\left(\theta\right)/\theta^{1-\chi}$, we get

$$\begin{aligned}\frac{\alpha\xi}{A^{\chi}}\int_{\Theta}\mu\left(\theta\right)\frac{\pi\left(\theta\right)}{\theta^{1-\chi}}d\theta \;=\;& \frac{\alpha\xi}{A^{\chi}}\overbrace{\left[\mu\left(\theta\right)\Pi\left(\theta\right)\right]\Big|_{\underline{\theta}}^{\overline{\theta}}}^{=0} \\ &+\frac{\alpha\xi}{A^{\chi}}\int_{\Theta}\Pi\left(\theta\right)\left[\frac{1}{\frac{\partial u(c)}{\partial c}}-\lambda(\theta_{-})-\gamma(\theta_{-})\frac{\frac{\partial f(\theta|\theta_{-})}{\partial\theta_{-}}}{f(\theta|\,\theta_{-})}\right]f(\theta|\,\theta_{-})d\theta\,.\end{aligned}$$

The first term on the right-hand side is equal to zero because of the boundary conditions for $\mu(\theta)$. We then note that the first-order condition (35) and equation (38) imply that

$$\gamma(\theta_-) = -\frac{\beta}{q}\frac{\mu(\theta_-)}{f(\theta_-|\,\theta_{--})} = -\frac{\beta}{q}\left(\frac{\tau^p_{l-}}{1-\tau^p_{l-}}\right)\frac{1}{\frac{\partial u(c_-)}{\partial c_-}}\frac{A^{\chi}\theta_-^{1-\chi}}{\alpha\xi}\;.$$

Furthermore, we can use the optimality condition (34) to replace $\lambda(\theta_-)$ with $\beta/\left[q\partial u\left(c\left(\theta_-\right)\right)/\partial c(\theta_-)\right]$, so that, after rearranging, (39) reads

$$\begin{aligned}&\int_{\Theta}\left(\frac{\tau^p_l}{1-\tau^p_l}\right)\frac{q}{\beta}\frac{\frac{\partial u(c_-)}{\partial c_-}}{\frac{\partial u(c)}{\partial c}}\pi(\theta)\,f(\theta|\,\theta_-)d\theta\\ =\;&\frac{\tau^p_{l-}}{1-\tau^p_{l-}}\theta_-^{1-\chi}\int_{\Theta}\Pi(\theta)\,\frac{\partial f(\theta|\,\theta_-)}{\partial\theta_-}d\theta+\frac{\alpha\xi}{A^{\chi}}\int_{\Theta}\Pi(\theta)\left[\frac{q}{\beta}\frac{\frac{\partial u(c_-)}{\partial c_-}}{\frac{\partial u(c)}{\partial c}}-1\right]f(\theta|\,\theta_-)d\theta\;.\end{aligned}$$

■

The labor wedge is similar to the one derived by Fahri and Werning (2013), Proposition 2. Notice that without persistence of θ, the first term on the right-hand side vanishes.

Specifying $\pi(\theta)=\theta^{-\chi}$ so that $\pi(\theta)/\theta^{1-\chi}=1/\theta$ and $\Pi(\theta)=\log(\theta)$, we get

$$\begin{aligned}&\int_{\Theta}\left(\frac{\tau^p_l}{1-\tau^p_l}\right)\frac{q}{\beta}\frac{\partial u(c_-)/\partial c_-}{\partial u(c)/\partial c}\theta^{-\chi}f(\theta|\,\theta_-)d\theta\\ =\;&\frac{\tau^p_{l-}}{1-\tau^p_{l-}}\theta_-^{1-\chi}\int_{\Theta}\log(\theta)\,\frac{\partial f(\theta|\,\theta_-)}{\partial\theta_-}d\theta\\ &+\frac{\alpha\xi}{A^{\chi}}\int_{\Theta}\log(\theta)\left[\frac{q}{\beta}\frac{\partial u(c_-)/\partial c_-}{\partial u(c)/\partial c}-1\right]f(\theta|\,\theta_-)d\theta\;.\end{aligned}$$

We can interpret $\frac{q}{\beta}\frac{\partial u(c_-)/\partial c_-}{\partial u(c)/\partial c}\theta^{-\chi}$ on the left-hand side as a change of measure, analogous to Farhi and Werning (2013), if productivity $A(\theta,h)$ is Cobb-Douglas in θ and h so that $\chi=0$. In this case, $E\left[\frac{q}{\beta}\frac{\partial u(c_-)/\partial c_-}{\partial u(c)/\partial c}\theta^{-\chi}\right]=E\left[\frac{q}{\beta}\frac{\partial u(c_-)/\partial c_-}{\partial u(c)/\partial c}\right]=1.$

Concerning the wedge for human capital accumulation, we find:

Proposition 7 *If types θ are persistent, and assumptions **[A1]** and **[A2]** hold, the human capital wedge can decomposed as*

$$\tau^p_h=\Delta^p_l+\Delta^p_b+\Delta^p_i$$

with

$$
\begin{aligned}
\Delta_l^p &\equiv \frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_\Theta \frac{\partial y'}{\partial h'} \tau_l^{p\prime} f(\theta'|\,\theta) d\theta', \\
\Delta_b^p &\equiv \frac{1}{\frac{\partial g(h',h)}{\partial h'}} \mathbb{E}\left[\beta \frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}} - q\right] \mathbb{E}\left[\frac{\partial y'}{\partial h'} - \frac{\partial g(h'',h')}{\partial h'}\right] \\
&\quad + \frac{\beta}{\frac{\partial g(h',h)}{\partial h'} \frac{\partial u(c)}{\partial c}} cov\left(\frac{\partial u\,(c')}{\partial c'}, \frac{\partial y'}{\partial h'} - \frac{\partial g(h'',h')}{\partial h'}\right), \\
\Delta_i^p &\equiv -\frac{q}{\frac{\partial u(c)}{\partial c} \frac{\partial g(h',h)}{\partial h'}} cov\left(\frac{1}{\frac{\partial u(c')}{\partial c'}}, \frac{\partial v(y',\theta',h')}{\partial h'}\right) \\
&\quad + \frac{\beta}{\frac{\partial u(c)}{\partial c} \frac{\partial g(h',h)}{\partial h'}} \frac{A^\chi}{\alpha\xi} \frac{\tau_l^p}{1-\tau_l^p} \theta^{1-\chi} cov\left(\frac{\frac{\partial f(\theta'|\theta)}{\partial \theta}}{f(\theta'|\,\theta)}, \frac{\partial v(y,\theta',h')}{\partial h'}\right).
\end{aligned}
$$

Proof. Adding the wedge for human capital, analogous to the definition in (12), on both sides of (37) and rearranging, we find that

$$
\tau_h^p = \Delta_l^p + \Delta_b^p + \Delta_i^p
$$

with

$$
\Delta_l^p \equiv \frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_\Theta \frac{\partial y'}{\partial h'} \tau_l^{p\prime} f(\theta'|\,\theta) d\theta',
$$

$$
\begin{aligned}
\Delta_b^p &\equiv \frac{1}{\frac{\partial g(h',h)}{\partial h'}} \int_\Theta \left(\beta \frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}} - q\right) \left(\frac{\partial y'}{\partial h'} - \frac{\partial g(h'',h')}{\partial h'}\right) f(\theta'|\,\theta) d\theta' \\
&= \frac{1}{\frac{\partial g(h',h)}{\partial h'}} \mathbb{E}\left[\beta \frac{\frac{\partial u(c')}{\partial c'}}{\frac{\partial u(c)}{\partial c}} - q\right] \mathbb{E}\left[\frac{\partial y'}{\partial h'} - \frac{\partial g(h'',h')}{\partial h'}\right] \\
&\quad + \frac{\beta}{\frac{\partial g(h',h)}{\partial h'} \frac{\partial u(c)}{\partial c}} cov\left(\frac{\partial u\,(c')}{\partial c'}, \frac{\partial y'}{\partial h'} - \frac{\partial g(h'',h')}{\partial h'}\right),
\end{aligned}
$$

$$
\Delta_i^p \equiv -\frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_\Theta \mu'(\theta') \frac{\partial^2 v(y',\theta',h')}{\partial \theta' \partial h'} d\theta'.
$$

While Δ_l^p and Δ_b^p are straightforward counterparts to the respective terms that apply if types are not persistent (see Δ_l and Δ_b in Proposition 3), developing Δ_i^p yields further insights. We elaborate on the term Δ_i^p integrating by parts:

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

The first term on the right-hand side is equal to zero because of the boundary conditions for $\mu'(\theta')$. Thus,

$$\begin{aligned}
&\Delta_i^p \\
=\;& \frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_\Theta \frac{\partial \mu'(\theta')}{\partial \theta'} \frac{\partial^2 v(y',\theta',h')}{\partial h'} d\theta' \\
=\;& -\frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_\Theta \left[\frac{1}{\partial u(c)/\partial c} - \lambda'(\theta) - \gamma'(\theta) \frac{\frac{\partial f(\theta'|\theta)}{\partial \theta}}{f(\theta'|\theta)} \right] \frac{\partial v(y',\theta',h')}{\partial h'} f(\theta'|\theta) d\theta'.
\end{aligned}$$

Since by (34) and (35),

$$\gamma'(\theta) = -\frac{\beta \mu(\theta)}{q f(\theta|\theta_-)}$$

and

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

we get

$$\begin{aligned}
&\Delta_i^p \\
=\;& -\frac{q}{\frac{\partial g(h',h)}{\partial h'}} \int_\Theta \left[\frac{1}{\frac{\partial u(c')}{\partial c'}} - \frac{\beta}{q \partial \frac{\partial u(c)}{\partial c}} + \frac{\beta \mu(\theta)}{q f(\theta|\theta_-)} \frac{\frac{\partial f(\theta'|\theta)}{\partial \theta}}{f(\theta'|\theta)} \right] \frac{\partial v(y',\theta',h')}{\partial h'} f(\theta'|\theta) d\theta' \\
=\;& -\frac{q}{\frac{\partial g(h',h)}{\partial h'} \frac{\partial u(c)}{\partial c}} \int_\Theta \left[\frac{\frac{\partial u(c)}{\partial c}}{\frac{\partial u(c')}{\partial c'}} - \frac{\beta}{q} + \frac{A^\chi \theta^{1-\chi}}{\alpha \xi} \frac{\beta}{q} \frac{\tau_l^p}{1-\tau_l^p} \frac{\frac{\partial f(\theta'|\theta)}{\partial \theta}}{f(\theta'|\theta)} \right] \frac{\partial v(y',\theta',h')}{\partial h'} f(\theta'|\theta) d\theta',
\end{aligned}$$

where the second equality follows from (38). In order to further simplify, note that, as in the case without persistent types, we have

$$\begin{aligned}
& \int_\Theta \left[\frac{1}{\partial u(c')/\partial c'} - \frac{\beta}{q \partial u(c)/\partial c} \right] \frac{\partial v(y',\theta',h')}{\partial h'} f(\theta'|\theta) d\theta' \\
=\;& \underbrace{\mathbb{E}\left[\frac{1}{\partial u(c')/\partial c'} - \frac{\beta}{q \partial u(c)/\partial c} \right]}_{=0} \mathbb{E}\left[\frac{\partial v(y',\theta',h')}{\partial h'} \right] \\
& + cov\left(\frac{1}{\partial u(c')/\partial c'} - \frac{\beta}{q \partial u(c)/\partial c}, \frac{\partial v(y',\theta',h')}{\partial h'} \right) \\
=\;& cov\left(\frac{1}{\frac{\partial u(c')}{\partial c'}}, \frac{\partial v(y',\theta',h')}{\partial h'} \right).
\end{aligned}$$

Moreover, the changes $\partial f(\theta'|\theta)/\partial \theta$ in the density have to sum to zero across all θ' so that

$$\mathbb{E}\left[\frac{\frac{\partial f(\theta'|\theta)}{\partial \theta}}{f(\theta'|\theta)} \right] = \int_\Theta \frac{\frac{\partial f(\theta'|\theta)}{\partial \theta}}{f(\theta'|\theta)} f(\theta'|\theta) d\theta' = \int_\Theta \frac{\partial f(\theta'|\theta)}{\partial \theta} d\theta' = 0.$$

It follows that

$$\begin{aligned}
&\int_\Theta \frac{\frac{\partial f(\theta'|\theta)}{\partial\theta}}{f(\theta'|\theta)} \frac{\partial v(y',\theta',h')}{\partial h'} f(\theta'|\theta) d\theta' \\
= \; & \underbrace{\mathbb{E}\left[\frac{\frac{\partial f(\theta'|\theta)}{\partial\theta}}{f(\theta'|\theta)}\right]}_{=0} \mathbb{E}\left[\frac{\frac{\partial v(y',\theta',h')}{\partial h'}}{f(\theta'|\theta)}\right] + cov\left(\frac{\frac{\partial f(\theta'|\theta)}{\partial\theta}}{f(\theta'|\theta)}, \frac{\partial v(y',\theta',h')}{\partial h'}\right) \\
= \; & cov\left(\frac{\frac{\partial f(\theta'|\theta)}{\partial\theta}}{f(\theta'|\theta)}, \frac{\partial v(y',\theta',h')}{\partial h'}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\Delta_i^p \\
= \; & -\frac{q}{\frac{\partial g(h',h)}{\partial h'}} cov\left(\frac{1}{\frac{\partial u(c')}{\partial c'}}, \frac{\partial v(y',\theta',h')}{\partial h'}\right) \\
& -\frac{\beta}{\frac{\partial u(c)}{\partial c}\frac{\partial g(h',h)}{\partial h'}} \frac{A^\chi}{\alpha\xi} \frac{\tau_l^p}{1-\tau_l^p} \theta^{1-\chi} cov\left(\frac{\frac{\partial f(\theta'|\theta)}{\partial\theta}}{f(\theta'|\theta)}, \frac{\partial v(y',\theta',h')}{\partial h'}\right).
\end{aligned}$$

■

Analogous to Proposition 3, Proposition 7 shows that the wedge for human capital is affected by the expected labor wedges in the next period in term Δ_l^p, the wedge for bequests in term Δ_b^p and the effect of human capital on the incentive-compatibility constraint in term Δ_i^p. Compared with the results for i.i.d types, the effect of human capital on the incentive-compatibility constraint in Δ_i^p also depends on the current labor wedge τ_l^p if ability types are persistent and $\gamma(\theta) > 0$. The sign of this additional effect depends on how the likelihood ratio $\frac{\partial f(\theta'|\theta)}{\partial\theta}/f(\theta'|\theta)$ covaries with the effect of human capital on the disutility of labor $\partial v(y',\theta',h')/\partial h'$ as θ' changes. We find:

Corollary 4 *Under assumptions **[A1]**, **[A1']** and **[A2]**, $\Delta_i^p < 0$ if $\chi \geq -\alpha$ and $\frac{\partial f(\theta'|\theta)}{\partial\theta}/f(\theta'|\theta)$ monotonically increases in θ'. The planner then has a motive to increase human capital accumulation in order to relax the incentive compatibility constraint.*

Proof. The proof follows directly from Remark 1. ■

It seems natural that $\frac{\partial f(\theta'|\theta)}{\partial\theta}/f(\theta'|\theta)$ increases in θ' since this implies that the planner is more likely to observe higher future output of dynasties that have high current ability. See, for example, the interpretation of the monotone likelihood ratio assumption in Rogerson (1985). With persistence of ability types, the planner thus has an additional incentive to subsidize education for reducing information rents of the future generation, and this incentive is stronger the larger is the current labor wedge τ_l^p.

Concerning the wedge for bequests, we impose assumption **[A1]**, use equation (33) and follow the steps of the derivations of the reciprocal Euler equation noting that $\mathbb{E}\left[\frac{\partial f(\theta'|\theta)}{\partial \theta}/f(\theta'|\theta)\right]=0$. This establishes that the wedge for bequests $\tau_b^p>0$ also in the case with persistent types. See also Stantcheva (2014).

References

[1] Abreu Dilip, David Pearce and Ennio Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," Econometrica, vol. 58(5), 1041-1063

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

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

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

[5] Chetty, Raj (2012): "Bounds on Elasticities with Optimization Frictions: a Synthesis of Micro and Macro Evidence on Labor Supply," Econometrica, vol. 80, 969–1018.

[6] Cunha, Flavio, James J. Heckman, Lance Lochner and Dimitriy V. Masterov (2006): "Interpreting the Evidence on Life Cycle Skill Formation," ch. 12, in: *Handbook of the Economics of Education*, Hanushek, Eric A. and Finis Welch (eds.), vol. 1, 697–812.

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

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

[9] Gelber, Alexander and Matthew Weinzierl (2014): "Equalizing Outcomes and Equalizing Opportunities: Optimal Taxation when Children's Abilities Depend on Parents' Resources," Working Paper No. 13-014, Harvard Business School.

[10] Golosov, Mikhail, Maxim Troshkin, Aleh Tsyvinski (2011): "Optimal Dynamic Taxes," NBER Working Paper No. 17642.

[11] Hanushek, Eric A. (2006): "School Resources," ch. 14, in: *Handbook of the Economics of Education*, Hanushek, Eric A. and Finis Welch (eds.), vol. 2, 865–908.

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

[13] Kapička, Marek (2013): "Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: a First-Order Approach," Review of Economic Studies, vol. 80, 1027–1054.

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

[15] Krueger, Dirk and Alexander Ludwig (2013): "Optimal Capital and Progressive Labor Income Taxation with Endogenous Schooling Decisions and Intergenerational Transfers," Manuscript, University of Pennsylvania.

[16] Lee, Sang Yoon (Tim) and Ananth Seshadri (2014): "On the Intergenerational Transmission of Economic Status," Manuscript, University of Mannheim.

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

[18] Pavan, Alessandro, Ilya Segal and Juuso Toikka (forthcoming): "Dynamic Mechanism Design: A Myersonian Approach," Econometrica.

[19] Rogerson, William P. (1985): "The First-Order Approach to Principal-Agent Problems," Econometrica, vol. 53, 1357–1367.

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