RUHR
ECONOMIC PAPERS

Christoph Braun

Optimal Taxation of Education with an Initial Endowment of Human Capital

Ruhr Graduate School in Economics
technische universität dortmund

#210

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, email: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Christoph M. Schmidt
RWI, Phone: +49 (0) 201/81 49-227, e-mail: christoph.schmidt@rwi-essen.de

Editorial Office

Joachim Schmidt
RWI, Phone: +49 (0) 201/81 49-292, e-mail: joachim.schmidt@rwi-essen.de

Ruhr Economic Papers #210

Responsible Editor: Wolfgang Leininger

ISSN 1864-4872 (online) – ISBN 978-3-86788-241-5

Ruhr Economic Papers #210

Christoph Braun

Optimal Taxation of Education with an Initial Endowment of Human Capital

technische universität dortmund

Ruhr Graduate School in Economics

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-241-5

Christoph Braun[1]

Optimal Taxation of Education with an Initial Endowment of Human Capital

Abstract

Bovenberg and Jacobs (2005) and Richter (2009) derive the education efficiency theorem: In a second-best optimum, the education decision is undistorted if the function expressing the stock of human capital features a constant elasticity with respect to education. I drop this assumption. The household inherits an initial stock of human capital, implying that the aforementioned elasticity is increasing. In a two-period Ramsey model of optimal taxation, I show that the education efficiency theorem does not hold. In a second-best optimum, the discounted marginal social return to education is smaller than the marginal social cost. The household overinvests in human capital relative to the first best. The government effectively subsidizes the return to education.

JEL Classification: H21, I28, J24

Keywords: Optimal taxation; human capital

October 2010

1 Ruhr Graduate School in Economics and TU Dortmund. *– I am grateful to Wolfram F. Richter, Spencer Bastani, and the audience in Dortmund and at the IEB summer school 2009 in Barcelona for helpful comments. I also thank the Gottfried Wilhelm Leibniz Scientific Community for financial support. The usual disclaimer applies.* *– All correspondence to Christoph Braun, Ruhr Graduate School in Economics, c/o TU Dortmund, Department of Economics, 44221 Dortmund, Germany, E-Mail: christoph.braun@tu-dortmund.de.*

1 Introduction

Bovenberg and Jacobs (2005) and Richter (2009) set up two-period models of the Mirrlees and of the Ramsey type and derive the so-called *education efficiency theorem*: In a second-best optimum, the education decision is undistorted. The before- and after-tax rates of return to education are equal. This result crucially depends on the way the accumulation of human capital is modeled, which is as follows. In the first period, the household spends time on education, which enters a human capital production function as the only input. The output increases the stock of human capital in the second period. Bovenberg and Jacobs (2005) and Richter (2009) assume that the stock of human capital in the second period equals only the output of the human capital production function, which is assumed to be isoelastic with respect to education.[1] This means that the function expressing the stock of human capital in the second period and the human capital production function are the same. A debatable implication of modeling the law of motion for human capital this way is that the stock of human capital in the second period is zero if the household does not spend any time on education in the first period. Then effective labor supply is zero, because it is modeled as the product of raw labor supply and the then existing stock of human capital. Consequently, the household does not earn any labor income. Put more briefly, if the household wants to reap benefits of human capital, it first has to spend time on education. Or, as an alternative interpretation of this implication, consider a two-period

[1] Jacobs and Bovenberg (2008) further analyze the human capital production function's properties and find that a constant elasticity is crucial for their results in Bovenberg and Jacobs (2005). Richter (2009) refers to the so-called power law of learning, a result from cognitive psychology that provides evidence in favor of a constant elasticity. See Ritter and Schooler (2001) for more details.

overlapping-generations (OLG) model. When young, the household accumulates human capital, which it uses when old. When old, it passes on the then existing stock of human capital to the newly born young household so that it can further increase the stock by spending time on education. But when a young household stops devoting time to education, the stock of human capital in the second period is zero. Consequently, the old household could not pass on human capital to the newly born young household.

By contrast, I assume that in the first period the household is endowed with an initial stock of human capital. The present paper then studies the effect that the initial stock of human capital has on optimal taxation. The stock of human capital in the second period is assumed to be the sum of the output of the isoelastic human capital production function, which takes education as the only input, and the initial stock of human capital net of depreciation. This implies that the elasticity of the function expressing the stock of human capital in the second period is increasing. With this specification, the stock of human capital in the second period does not drop to zero even when the household stops spending time on education in the first period, or, referring back to the OLG interpretation, the old household can then pass on human capital even when it may not have spent time on education when young. Then the education efficiency theorem no longer holds. In a second-best optimum, the discounted marginal social return to education is smaller than the marginal social cost. The household overinvests in human capital relative to the first best. As a result, the government effectively subsidizes the return to education.

The general-equilibrium model used here comprises a single household, a firm, and a government. The household lives for two periods, in which it faces a consumption-labor-leisure choice. In the first period, it chooses how much time to devote to work and education. In the second period, it only decides how much to work. Time spent on education is transformed into human capital by means of a human capital production function. The household combines its raw labor supply with the then existing stock of human capital, giving the effective labor supply. It chooses to lend capital to a firm, which takes it as an input, jointly with the effective labor supply, and pays a return. The firm produces a single consumption good. Time spent on education brings about disutility and comes at the cost of forgone earnings and some direct costs such as tuition fees. All actions of the household are assumed to be fully observable. The government levies linear taxes on the household's income from work and saving to finance an exogenously given stream of expenditures. Furthermore, it may choose to subsidize the direct cost of education. The question then is how to optimally choose linear taxes and the subsidy to maximize the household's utility given exogenous government expenditures and subject to the household's competitive equilibrium behavior.

2 The Model

2.1 Household's Problem

The household solves the following maximization problem:

$$\begin{aligned}\mathcal{L} &= U(C_0, L_0 + E) + \beta U(C_1, L_1) \\ &+ \lambda_0\Big(\omega_0 L_0 H_0 + R_0^\tau K_0 - C_0 - K_1 - \varphi E\Big) + \beta\lambda_1\Big(\omega_1 L_1 H_1 + R_1^\tau K_1 - C_1\Big) \\ &\qquad + \mu\Big(G(E) + (1-\delta_H)H_0 - H_1\Big). \qquad (1)\end{aligned}$$

The household's utility function is strictly increasing in consumption, C_t, and strictly decreasing in the nonleisure times $L_0 + E$ and L_1. It is strictly concave in both arguments and time-separable.

Savings serve as a means to smooth consumption over time. They pay the net rate of return $R_t^\tau \equiv (1-\tau_t^K)r_t + 1 - \delta_K$, where τ_t^K is a linear tax on the gross rate of return r_t, and δ_K is the rate at which the stock of capital K_t depreciates. Raw labor supply L_t is combined with the stock of human capital H_t accumulated so far. Effective labor supply $L_t H_t$ earns the net wage rate $\omega_t \equiv (1-\tau_t^L)w_t$, where τ_t^L is a linear tax on the gross wage rate w_t. Let $\varphi \equiv (1-\tau^H)f$ be the direct cost of education net of the subsidy τ^H, where f is an exogenous (fee) parameter. The endowments of the initial stocks of human capital, H_0, and capital, K_0, are given. β is the private discount factor.

The law of motion

$$H_1(E) = G(E) + (1-\delta_H)H_0 \tag{2}$$

governs the evolution of the stock of human capital, which depreciates at the rate $\delta_H \leq 1$. Here G is the human capital production function, which takes time E as its only input factor. It is isoelastic:

$$G(E) = aE^\gamma \tag{3}$$

with $0 < \gamma < 1$. The coefficient $a > 0$ is a shift parameter. (2) and (3) imply that the elasticity η of the function of the stock of human capital $H_1(E)$ is strictly increasing as long as the initial stock of human capital does not fully depreciate.[2] By setting $H_0 = 0$ or $\delta_H = 1$, one obtains the model

[2] *Proof.* $\frac{d}{dE}\eta \equiv \frac{d}{dE}\frac{EH_1'(E)}{H_1(E)} = \gamma G'(E)\frac{H_1(E)-G(E)}{H_1(E)^2} > 0$ for $\delta_H < 1$. $\square$

underlying the analysis in Bovenberg and Jacobs (2005) or Richter (2009). Then the elasticity of $H_1(E)$ equals γ.

To have a well-behaved problem, it does not suffice to apply the Inada conditions. An analysis of the second-order conditions, which is done in appendix A, reveals that moreover one has to assume that the utility function is sufficiently concave to compensate for the lack of concavity of the law of motion (2) for human capital. Put formally, the requirement says that $\gamma < v_1/(1+v_1)$, where γ is the elasticity of the human capital production function (3), and $v_1 = L_1 U_{L_1L_1}/U_{L_1}$, which captures the concavity of the utility function with respect to second-period labor supply. This condition will show up again when labor taxation is analyzed in section 2.7.3.

Let U_{C_t} and U_{L_t} denote the partial derivatives with respect to consumption and nonleisure time, taking the corresponding period t variables as arguments. Maximization over consumption, time spent on working, and investments in human and physical capital yields the following first-order conditions:

$$C_t : U_{C_t} = \lambda_t,\ t = 0, 1, \tag{4}$$

$$L_t :\ -U_{L_t} = \omega_t H_t \lambda_t,\ t = 0, 1, \tag{5}$$

$$E :\ -U_{L_0} + \lambda_0 \varphi = \mu G'(E), \tag{6}$$

$$H_1 : \lambda_1 \beta \omega_1 L_1 = \mu, \tag{7}$$

$$K_1 : \lambda_0 = \lambda_1 \beta R_1^{\tau}. \tag{8}$$

By eliminating μ and using all first-order conditions, the following optimality condition results:

$$\frac{\omega_1 L_1 G'(E)}{R_1^{\tau}} = \varphi + \omega_0 H_0. \tag{9}$$

The household chooses education up to the point where the discounted marginal (private) return $\omega_1 L_1 G'(E)/R_1^{\tau}$ equals the marginal (private) cost $\varphi + \omega_0 H_0$, which is sum of the direct cost and the forgone earnings.

2.2 The Government

The government uses linear taxes to finance an exogenously given stream of government expenditures $\{g_t\}_{t=0}^{1}$. Its budget constraints are

$$g_0 + \tau^H f E = \tau_0^K r_0 K_0 + \tau_0^L w_0 L_0 H_0, \tag{10}$$

$$g_1 = \tau_1^K r_1 K_1 + \tau_1^L w_1 L_1 H_1. \tag{11}$$

2.3 Firm's Problem

The stock of physical capital K_t and the household's effective labor supply, $Z_t \equiv L_t H_t$, enter the firm's constant-returns-to-scale production function $F(K_t, Z_t)$. Factors are paid their marginal products:

$$F_{K_t} \equiv \frac{\partial}{\partial K_t} F(K_t, Z_t) = r_t, \; t = 0, 1, \tag{12}$$

$$F_{Z_t} \equiv \frac{\partial}{\partial Z_t} F(K_t, Z_t) = w_t, \; t = 0, 1. \tag{13}$$

2.4 Competitive Equilibrium

A competitive equilibrium consists of a feasible allocation

$$\Big\{\{C_t, L_t, K_t, H_t\}_{t=0}^{1}, E\Big\},$$

a price system

$$\{w_t, r_t\}_{t=0}^{1},$$

a government policy

$$\Big\{\{g_t, \tau_t^K, \tau_t^L\}_{t=0}^{1}, \tau^H\Big\},$$

an exogenously given direct cost of education f, and initial stocks of human and physical capital, H_0 and K_0, respectively. The feasible allocation and the price system solve the household's and firm's problems. The government policy satisfies the budget constraints (10) and (11).

2.5 First-Best Solution

Studying the first-best problem serves to establish a benchmark case. The planner chooses consumption, investments in physical and human capital, and the allocation of time to solve the following maximization problem:

$$\begin{aligned}\mathcal{L} = {} & U(C_0, L_0 + E) + \beta U(C_1, L_1) \\ & + \theta_0 \Big(F(K_0, L_0 H_0) + (1 - \delta_K) K_0 - C_0 - K_1 - fE - g_0 \Big) && (14) \\ & + \theta_1 \beta \Big(F(K_1, L_1 H_1) + (1 - \delta_K) K_1 - C_1 - g_1 \Big) && (15) \\ & + \mu \Big(G(E) + (1 - \delta_H) H_0 - H_1 \Big).\end{aligned}$$

He maximizes the household's discounted sum of utilities subject to the per-period resource constraints (14) and (15) and the law of motion for human capital.

The first-order conditions are

$$
\begin{aligned}
C_t &: \; U_{C_t} = \theta_t, \; t = 0, 1, && (16)\\
L_t &: \; -U_{L_t} = \theta_t F_{Z_t} H_t, \; t = 0, 1, && (17)\\
E &: \; \mu G'(E) = -U_{L_0} + \theta_0 f, && (18)\\
K_1 &: \; \theta_0 = \theta_1 \beta \big(F_{K_1} + 1 - \delta_K\big) \equiv \theta_1 \beta R_1^s, && (19)\\
H_1 &: \; \theta_1 \beta F_{Z_1} L_1 = \mu. && (20)
\end{aligned}
$$

Proposition 1. *The discounted marginal social return to education equals the marginal social cost:*

$$\frac{F_{Z_1} L_1 G'(E)}{R_1^s} = f + F_{Z_0} H_0. \tag{21}$$

Proof. Eliminate θ_0, μ, and U_{L_0} in the condition (18) using (17), (19), and (20). □

The social planner chooses education up to the point where the discounted marginal (social) return $F_{Z_1} L_1 G'(E) / R_1^s$ equals the marginal (social) cost $f + F_{Z_0} H_0$, which is sum of the direct cost and the loss in marginal productivity of first period's labor supply. Proposition 1 therefore suggests the following definition to gauge education efficiency.

Definition 1. *Education efficiency is achieved if the discounted marginal social return to education equals the marginal social cost, which is the direct cost of education plus the loss in marginal productivity of the first period's labor supply. In the first best, there is no wedge between the discounted marginal social return and the marginal social cost of education.*

The efficiency condition (21) can be further used to assess under which circumstances a competitive equilibrium implies education efficiency in the sense of Definition 1. The wedge between the discounted marginal social return and the marginal social cost of education can be manipulated as follows:

$$\begin{aligned}\Delta &\equiv \frac{F_{Z_1} L_1 G'(E)}{R_1^s} - \left(f + F_{Z_0} H_0\right) \\ &= \frac{R_1^\tau\left(\varphi + \omega_0 H_0\right)}{R_1^s}\left[\frac{F_{Z_1}}{\omega_1} - \frac{R_1^s\left(f + F_{Z_0} H_0\right)}{R_1^\tau(\varphi + \omega_0 H_0)}\right].\end{aligned} \tag{22}$$

The last equality follows from the household's optimality condition (9). Education efficiency holds if and only if the bracketed factor vanishes. Therefore,

$$\frac{F_{Z_1}}{R_1^s\left(f + F_{Z_0} H_0\right)} = \frac{\omega_1}{R_1^\tau\left(\varphi + \omega_0 H_0\right)}. \tag{23}$$

Put verbally, if and only if before- and after-tax rates of return are equal, education efficiency prevails in a competitive equilibrium. The wedge Δ is positive (negative) if and only if education is effectively taxed (subsidized). Richter (2009) uses the condition (23) to assess education efficiency.

2.6 Second-Best Solution

The Ramsey problem is to choose a government policy that maximizes the household's utility subject to the household's and the firm's competitive equilibrium behavior, given initial stocks of human and physical capital, H_0 and K_0, and direct cost of education f. The primal approach is adopted to study the Ramsey problem of optimal taxation (Atkinson and Stiglitz (1980), Chari and Kehoe (1999)). The difference to the dual approach is how it incorporates the household's competitive equilibrium behavior.

The household's first-order conditions serve to eliminate all prices and taxes in the intertemporal budget constraint. As a result, this constraint then fully captures how the household behaves in a competitive equilibrium. Given the allocation, the first-order conditions yield the prices and taxes that implement the second-best outcome as a competitive equilibrium. By contrast, the dual approach includes all constraints separately, which requires optimizing over the allocations and prices.

To derive the so-called implementability constraint, first the intertemporal budget results after combining the per-period budget constraints from the household's problem (1) by eliminating K_1 and using (9) to eliminate direct cost φE:

$$R_0^\tau K_0 + \omega_0 H_0 (L_0 + E) + \frac{1}{R_1^\tau} \omega_1 L_1 H_1 \left(1 - \eta\right) = C_0 + \frac{1}{R_1^\tau} C_1 \tag{24}$$

with

$$\eta \equiv \eta(E, H_1) \equiv \frac{H_1'(E)E}{H_1(E)} = \frac{G'(E)E}{H_1(E)},$$

which is the nondecreasing[3] elasticity of the function expressing the stock of human capital in the second period. The LHS of (24) is the household's income side. $R_0^\tau K_0$ is the value of the initial stock of physical capital. The RHS is the expenditure side.

Using the household's first-order conditions (4), (5), and (8), the intertemporal budget constraint (24) can be written as

$$A = U_{C_0} C_0 + \beta U_{C_1} C_1 + U_{L_0} (L_0 + E) + \beta U_{L_1} L_1 \left(1 - \eta\right) \tag{25}$$

with

[3] See footnote 2.

$$A \equiv U_{C_0} R_0^{\tau} K_0, \tag{26}$$

which is a function of the endogenous variables C_0, L_0, E, and τ_0^K, and of the exogenous variables K_0 and H_0.

The allocations that the household's problem imply for a given government policy satisfy the implementability constraint (25) and the per-period resource constraints (14) and (15) (see Proposition 1 in Chari and Kehoe (1999)).

The government commits to a specific policy chosen at the outset of period 0, meaning that it does not reoptimize during the course of time.

The Ramsey problem reads

$$\begin{aligned}\mathcal{L} = U(C_0, L_0 + E) + \beta U(C_1, L_1) \\ + \theta_0 \Big(F(K_0, L_0 H_0) + (1 - \delta_K) K_0 - C_0 - K_1 - fE - g_0 \Big) \\ + \beta\theta_1 \Big(F(K_1, L_1 H_1) + (1 - \delta_K) K_1 - C_1 - g_1 \Big) \\ + \mu \Big(G(E) + (1 - \delta_H) H_0 - H_1 \Big) \\ + \phi \Big(U_{C_0} C_0 + \beta U_{C_1} C_1 + U_{L_0} (L_0 + E) + \beta U_{L_1} L_1 \left(1 - \eta\right) - A \Big).\end{aligned}$$

The following assumption simplifies the derivation of the first-order conditions.

Assumption 1. *The utility function U is additively separable in consumption and nonleisure, that is,* $U_{C_t L_t} = 0, \; t = 0, 1.$

The first-order conditions for the Ramsey problem are

$$C_0 : \; U_{C_0} - \theta_0 + \phi \Big(U_{C_0 C_0} C_0 + U_{C_0} - A_{C_0} \Big) = 0, \tag{27}$$

$$C_1: \ U_{C_1} - \theta_1 + \phi\Big(U_{C_1C_1}C_1 + U_{C_1}\Big) = 0, \tag{28}$$

$$L_0: \ U_{L_0} + \theta_0 F_{Z_0} H_0 + \phi\Big(U_{L_0L_0}(L_0 + E) + U_{L_0} - A_{L_0}\Big) = 0, \tag{29}$$

$$L_1: \ U_{L_1} + \theta_1 F_{Z_1} H_1 + \phi\Big(U_{L_1L_1}L_1 + U_{L_1}\Big)(1-\eta) = 0, \tag{30}$$

$$\begin{aligned} E: \ & U_{L_0} - \theta_0\varphi + \mu G'(E) \\ & \quad + \phi\left(U_{L_0L_0}(L_0 + E) + U_{L_0} - \beta U_{L_1} L_1 \frac{d\eta}{dE}\right) = 0, \end{aligned} \tag{31}$$

$$K_1: \ -\theta_0 + \beta\theta_1\left(F_{K_1} + 1 - \delta_K\right) = 0, \tag{32}$$

$$H_1: \ \beta\theta_1 F_{Z_1} L_1 - \mu - \phi\beta U_{L_1} L_1 \frac{d\eta}{dH_1} = 0. \tag{33}$$

Maximizing over τ_0^K would be the same as taxing away the return to the initial stock of capital, which essentially is a lump-sum tax.[4] Assuming $\tau_0^K = 0$ rules out this form of taxation.

2.7 Results

2.7.1 Taxation of Physical Capital

To study the case of taxation of physical capital, the following assumption limits the analysis to a specific type of utility functions.

[4] To see this point, maximize the Lagrangian over τ_0^K:

$$\frac{\partial \mathcal{L}}{\partial \tau_0^K} = \phi U_{C_0} F_{K_0} K_0$$

Introducing a lump-sum tax, namely τ_0^K, enhances welfare, as less distortionary taxation is necessary. ϕ measures the cost of using distortionary taxation. The other three factors are positive. Therefore, $\phi > 0$.

Optimally, τ_0^K should be chosen such that all government expenditures could be financed. Then we would have $\phi = 0$ and the present problem coincides with the first-best problem. This renders the whole analysis uninteresting. See also Jones, Manuelli, and Rossi (1997, p. 111).

Assumption 2. *The instantaneous utility function shall have the following form:*

$$U(C_t, \cdot) = \begin{cases} \frac{C_t^{1-\sigma}-1}{1-\sigma} - V(\cdot),\ t = 0, 1;\ 0 \leq \sigma \neq 1 \\ \ln C_t - V(\cdot),\ t = 0, 1;\ \sigma = 1 \end{cases}$$

V is strictly increasing and strictly convex. It is a function of $L_0 + E$ and L_1, respectively. $1/\sigma$ is the intertemporal elasticity of substitution in consumption.

Proposition 2. *Given Assumption 2 with $\sigma > 0$, if $R_0^\tau K_0 > 0$, then $\tau_1^K > 0$.*

Proof. Combine the conditions (27), (28), and (32):

$$\frac{U_{C_0}}{\beta U_{C_1}} \frac{1 + \phi\,(1-\sigma) - \phi \frac{U_{C_0 C_0}}{U_{C_0}} R_0^\tau K_0}{1 + \phi(1-\sigma)} = R_1^s.$$

To determine τ_1^K, use the household's conditions (4) and (8):

$$1 < \frac{1 + \phi\,(1-\sigma) - \phi \frac{U_{C_0 C_0}}{U_{C_0}} R_0^\tau K_0}{1 + \phi(1-\sigma)} = \frac{R_1^s}{R_1^\tau} = 1 + \frac{\tau_1^K r_1}{R_1^\tau}. \tag{34}$$

The proposed result follows immediately. □

Assumption 2 is necessary because it allows one to compare the denominator and numerator in (34), which would not be possible if the coefficient σ were not constant.

Proposition 2 is a well-known result in macroeconomics.[5] Taxation of the return to physical capital in the first period was ruled out by assumption, and therefore the government was not able to extract the profit coming from the initial stock of physical capital. In period 1, the positive tax on capital income is due to this initial stock. One may view this capital tax

[5] See, for instance, Proposition 7 in Chari and Kehoe (1999).

as an attempt to take away part of the return to capital, which was ruled out in the first period.

The proof of Proposition 2 highlights also that taxation of physical capital income depends on the household's preferences for consumption. To emphasize this point, suppose that the household's utility from consumption is linear, $\sigma = 0$. Then $U_{C_0C_0} = 0$, and $\tau_1^K = 0$ results.

The preceding analysis furthermore shows that the results are derived despite and not because of the presence of human capital in the model. This points out that taxes on the return to physical capital are not a vehicle to provide education incentives. As the next section will show, the wedge between the discounted marginal social return and the marginal social cost of education does not vanish with the optimal capital tax rate.

2.7.2 Taxation of Human Capital

Proposition 3. *The discounted marginal social return to education is smaller than the marginal social cost:*

$$\frac{F_{Z_1} L_1 G'(E)}{R_1^s} < f + F_{Z_0} H_0. \tag{35}$$

Proof. The first-order conditions (29) and (31) imply

$$\begin{aligned} &\beta \frac{\theta_1}{\theta_0} F_{Z_1} L_1 G'(E) - \left(f + F_{Z_0} H_0\right) \\ &\qquad = -\frac{\phi}{\theta_0} \left\{ \beta L_1 U_{L_1} \left(-G'(E) \frac{d\eta}{dH_1} - \frac{d\eta}{dE} \right) + A_{L_0} \right\}. \end{aligned} \tag{36}$$

By (32), $\beta\theta_1/\theta_0$ equals the social discount factor $1/R_1^s$. As a result, the LHS of (36) is the wedge Δ as defined by (22).

To prove the inequality, one has to determine the sign of the factor in curly brackets. The first term in it can be rearranged using the law of motion (2) for human capital, the specific functional form (3) of G, and the household's optimality condition (9). Therefore,

$$\Delta = -\frac{\phi}{\theta_0}\left\{\frac{\gamma}{H_1}\underbrace{(1-\delta_H)H_0}_{\#1}(\varphi+\omega_0)U_{C_0} + \underbrace{A_{L_0}}_{\#2}\right\} < 0.$$

From the definition (26), $A_{L_0} = U_{C_0}F_{K_0Z_0}H_0K_0$.[6] As long as $H_0 > 0$, the factor in the curly brackets is positive. It further increases as the human capital depreciation rate δ_H decreases. ϕ is positive for the reason explained above; see footnote 4, p. 15. □

Corollary 1. *In the second-best optimum, the household overinvests in human capital relative to the first best.*

Proof. One can show that the discounted marginal return to education is a decreasing function of E, ceteris paribus. The marginal cost is constant. As a result, the household overinvests in human capital relative to the first best. □

The results qualify the education efficiency theorem and show under which circumstances it does not hold. To begin, Proposition 3 holds in any case if there is an initial stock of human capital, $H_0 > 0$. Then at least #2 does not vanish. The source of distortion is the term A_{L_0}, which is due to the endogeneity of the first-period interest rate F_{K_0}. It is an initial endowment effect similar to the one discussed in the context of physical capital taxation. In Richter (2009) and Bovenberg and Jacobs (2005) this effect is not present, because they use a partial equilibrium analysis in which the interest rate is fixed, which means $F_{K_0Z_0} = 0$ and $A_{L_0} = 0$ results.

[6] Recall that $\tau_0^K = 0$ was assumed.

Regarding #1, if $H_1 \equiv G(E) = aE^{\gamma}$, which follows from setting $\delta_H = 1$ or $H_0 = 0$ in the law of motion (2) for human capital, the elasticity of the function $H_1(E)$ is constant, that is, $\frac{d\eta}{dE} = \frac{d\eta}{dH_1} = 0$. This means the positive product $(1 - \delta_H)H_0$ is one source of distortion, because as long as the initial stock of human capital does not fully depreciate, the elasticity η is increasing. This is the essence of Remark 2 in Richter (2009).

Both effects #1 and #2 would vanish with $H_0 = 0$, and the education efficiency theorem would result. This case corresponds to ruling out labor supply in the first period, as is done, for instance, in Jacobs and Bovenberg (2009).

The preceding results allow one to show that education is effectively subsidized. Use (35) and (9) to derive

$$\frac{F_{Z_1}}{R_1^s(f + w_0 H_0)} < \frac{\omega_1}{R_1^\tau(\varphi + \omega_0 H_0)} \Leftrightarrow \Delta < 0. \tag{37}$$

The before-tax rate of return to education is smaller than the after-tax rate of return, which means that the wedge Δ between the discounted marginal social return and the the marginal social cost is negative. Therefore, the following proposition results:

Proposition 4. *Education is effectively subsidized relative to the first best. The private rate of return to education is larger than the social rate of return.*

Physical and human capital are two assets, which the household can hold to smooth consumption over time. Above, on page 16, it is explained that it is optimal to tax the return to physical capital in the second period. Turning to human capital, the household disposes of an initial stock of human capital H_0. The return to it can only be taxed in a distorting way, because the tax rate τ_0^L does not have the characteristic of a lump-sum tax.

Consequently, this tax is an imperfect instrument, in the sense that it distorts the labor decision, to extract the return to the initial stock of human capital. For this reason, in the second period, when the household reaps the fruits of education, I therefore would have expected the government to at least partly skim off the additional return that could be attributed to the initial stock of human capital. The striking result, however, is that the contrary is true. The government should subsidize the accumulation of human capital. To sum up this point, an initial stock of physical capital implies that it is optimal to tax physical capital, whereas an initial stock of human capital calls for a subsidization of human capital. The intuition for why it is optimal to subsidize human capital may be the following. Labor taxation exerts a depressing effect on the accumulation of human capital. To counter this, a subsidy is helpful.

One may view the above result in a different light and interpret the model as the steady state of an OLG model as in Nielsen and Sørensen (1997). Then the OLG interpretation of the present model is in line with Propositions 2 and 3 in Richter and Braun (2010). The first result states that if the function $H_1(E)$ is isoelastic, education will remain undistorted. This case corresponds to setting $\delta_H = 1$ and thereby implicitly assuming that the young household does not inherit any stock of human capital from the old household. Term #1 vanishes. Proposition 3 states that if the elasticity of $H_1(E)$ is increasing, which is the case when $\delta_H < 1$, education will be subsidized relative to the first best. Also, the strength of the positive distortion depends on the cost resulting from the unavailability of lump-sum taxes, as captured by the Lagrange multiplier ϕ. Term #2 is not present in a steady state, because the initial endowment effect A_{L_0} only occurs in the first period and not later on.

2.7.3 Taxation of Labor

Proposition 5. *Labor tax rates are given by*

$$\frac{\tau_0^L}{1-\tau_0^L} = -\frac{\phi}{1+\phi}\left(v_0 + \frac{A_{L_0} + A_{C_0}F_{Z_0}H_0 + F_{Z_0}H_0U_{C_0C_0}C_0}{-U_{L_0}}\right), \tag{38}$$

$$\frac{\tau_1^L}{1-\tau_1^L} = -\frac{\phi}{1+\phi}\left((1-\eta)\,v_1 - \eta + \frac{U_{C_1C_1}}{-U_{L_1}}C_1F_{Z_1}H_1\right) \tag{39}$$

with

$$v_0 = \frac{(L_0+E)U_{L_0L_0}}{U_{L_0}} \text{ and } v_1 = \frac{L_1U_{L_1L_1}}{U_{L_1}}$$

denoting the reciprocals of the elasticities of nonleisure in periods 0 *and* 1 *in Frisch's sense.*[7]

Proof. Combine the first-order conditions (27) and (29), and (28) and (30). □

τ_0^L depends on initial endowment effects and the household's preferences for consumption. τ_1^L is affected by the effect of human capital, which is captured by the elasticity η.

The following assumption helps to gain further insight into (38) and (39).

Assumption 3.
1. *Interpret the above model as the steady state of an overlapping-generations model.*
2. *The utility function U shall be linear in consumption.*
3. *Human capital fully depreciates ($\delta_H = 1$), or the initial stock of human capital is zero ($H_0 = 0$).*

[7] L_t is implicitly defined by (5). Differentiating this condition with respect to, say, ω_0, holding λ_0 constant, yields $1/v_0 = U_{L_0}/[(L_0+E)U_{L_0L_0}]$. See Cahuc and Zylbergerg (2004, p. 20) for further details.

The first assumption implies that the initial endowment effects A_{L_0} and A_{C_0} are not present. $U_{C_tC_t} = 0$ follows from the second assumption, which implies that savings are not taxed in the second period: $\tau_1^K = 0$. The third assumption implies that the elasticity of the function expressing human capital in the second period equals the human capital production function's elasticity: $\eta = \gamma$.

Division of (39) by (38) then yields

$$\frac{\frac{\tau_1^L}{1-\tau_1^L}}{\frac{\tau_0^L}{1-\tau_0^L}} = \frac{(1-\gamma)v_1 - \gamma}{v_0}. \tag{40}$$

(40) is the analogue to equation 13 in Richter (2009), who terms it an extension of the inverse elasticity rule to cope with endogenous education.[8] He, however, assumed a utility function of the form $U = Z(C_0, C_1) - V(L_0 + E) - V(L_1)$, with the function Z being linear homogeneous. Because the utility function used here is time-separable in consumption, linear homogeneity means that utility is linear.

To have positive tax rates on labor income when young and old, the numerator in (40) has to be positive: $\gamma < v_1/(1+v_1)$. This inequality emerged from the analysis of the second-order conditions; see appendix A. It is a sufficient condition that must hold to have a well-behaved problem and moreover ensures that the tax rates are positive.

[8] Another insignificant difference is that Richter (2009) uses exclusive tax rates, whereas inclusive tax rates are used here.

$$1 - \tau = \frac{1}{1+\tau'}$$

is the formula for converting from a tax-inclusive basis to a tax-exclusive basis (Atkinson and Stiglitz, 1980, p. 70).

3 Conclusion

This paper has reassessed the models by Bovenberg and Jacobs (2005) and Richter (2009). Their and my papers tackle the same set of questions, use different approaches, but in the end come to similar conclusions. First, I demonstrated that the question of how to tax the return to physical capital is not affected by the accumulation of human capital. Taxing the return to physical capital investments is not a means to yield efficient investments in human capital. Then I showed that an increasing elasticity of the function expressing the stock of human capital in the second period implies subsidizing the return to education. An increasing elasticity arises if the household is endowed with an initial stock of human capital that does not fully depreciate.

The existing literature on Ramsey models of optimal taxation (see Atkeson, Chari, and Kehoe (1999) among others) and this paper both come to the conclusion that the question of whether to tax the return to capital or not depends on individual consumption preferences. In the second period this result is special, because one has to allow for an initial endowment effect, which is due to the initial stock of capital. The issue of capital taxation is independent of whether the model features human capital accumulation or not. To sum up this point, taxing capital income in the second period is optimal, but for other reasons than achieving efficient investments in human capital.

Further research could discuss the present model in an infinite-horizon setup, as it is done by Jones, Manuelli, and Rossi (1997). The major difference is that their human capital production function exhibits constant returns to scale with respect to stock variables that enter as means of production. By this specification they model human capital very symmetrically to physical capital and show that the return to education should remain untaxed in a steady state. The natural question then arises what exactly is the difference between human and physical capital. This paper works with a human capital production function that does not include the stock of human capital as a production factor. Then a model of optimal taxation could answer the question of how to tax the return to education if the human capital production function does not exhibit the restrictive properties as in Jones, Manuelli, and Rossi (1997). See also Ljungqvist and Sargent (2004, p. 534,) for a short discussion of this point.

References

ATKESON, A., V. V. CHARI, AND P. J. KEHOE (1999): "Taxing capital income: a bad idea," *Quarterly Review*, 23(3), 3–17.

ATKINSON, A. B., AND J. E. STIGLITZ (1980): *Lectures on Public Economics*. McGraw-Hill Inc.

BOVENBERG, A. L., AND B. JACOBS (2005): "Redistribution and education subsidies are Siamese twins," *Journal of Public Economics*, 89(11-12), 2005–2035.

CAHUC, P., AND A. ZYLBERGERG (2004): *Labor Economics*. MIT.

CHARI, V. V., AND P. J. KEHOE (1999): "Optimal fiscal and monetary policy," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1 of *Handbook of Macroeconomics*, chap. 26, pp. 1671–1745. Elsevier.

JACOBS, B., AND A. L. BOVENBERG (2008): "Optimal Taxation of Human Capital and the Earnings Function," CESifo Working Paper Series CESifo Working Paper No. 2250, CESifo GmbH.

——— (2009): "Human Capital and Optimal Positive Taxation of Capital Income," *International Tax and Public Finance*, forthcoming.

JONES, L. E., R. E. MANUELLI, AND P. E. ROSSI (1997): "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 73(1), 93–117.

LJUNGQVIST, L., AND T. J. SARGENT (2004): *Recursive Marcoeconomic Theory*. MIT Press, 2nd edn.

NIELSEN, S. B., AND P. B. SØRENSEN (1997): "On the optimality of the Nordic system of dual income taxation," *Journal of Public Economics*, 63(3), 311–329.

RICHTER, W. F. (2009): "Taxing Education in Ramsey's Tradition," *Journal of Public Economics*, 93(11-12), 1254–1260.

RICHTER, W. F., AND C. BRAUN (2010): "Efficient Subsidization of Human Capital Accumulation with Overlapping Generations and Endogenous Growth," CESifo Working Paper Series CESifo Working Paper No. 2982, CESifo Group Munich.

RITTER, F. E., AND L. J. SCHOOLER (2001): "The learning curve," in *International Encyclopedia of the Social and Behavioral Sciences*, ed. by P. B. W. Kintch, N. Smelser, pp. 8602–8605. Pergamon.

A Second-order Conditions

To study the conditions that must hold to have a well-behaved problem with an interior solution, the original problem (1) is written as one with only a single constraint. This is done by deriving the intertemporal budget constraint by substituting out the capital K_1, and then replacing the stock of human capital H_1 with the law of motion (2) for human capital. The Lagrangian therefore reads

$$\begin{aligned}&\mathcal{L} = U(C_0, L_0 + E) + \beta U(C_1, L_1)\\ &+ \lambda\left(R_0^\tau K_0 + \omega_0 L_0 H_0 + \frac{1}{R_1^\tau}\omega_1 L_1\big(G(E) + 1 - \delta_H\big) - C_0 - \frac{1}{R_1^\tau}C_1 - \varphi E\right).\end{aligned}$$

A sufficient condition for the solution to solve the constrained maximization problem is that the bordered Hessian of the Lagrangian satisfies the condition that the last four leading principal minors alternate in sign, the sign of the first one being positive. The bordered Hessian reads

$$H = \left(\begin{array}{c:ccccc} 0 & -1 & \omega_0 H_0 & \frac{1}{R_1^\tau}\omega_1 L_1 G' - \varphi & -\frac{1}{R_1^\tau} & \frac{1}{R_1^\tau}\omega_1 G \\ \hdashline -1 & U_{C_0C_0} & 0 & 0 & 0 & 0 \\ \omega_0 H_0 & 0 & U_{L_0L_0} & U_{L_0L_0} & 0 & 0 \\ \frac{1}{R_1^\tau}\omega_1 L_1 G' - \varphi & 0 & U_{L_0L_0} & U_{L_0L_0} + \lambda \frac{1}{R_1^\tau}\omega_1 L_1 G'' & 0 & \lambda \frac{1}{R_1^\tau}\omega_1 G' \\ -\frac{1}{R_1^\tau} & 0 & 0 & 0 & \beta U_{C_1C_1} & 0 \\ \frac{1}{R_1^\tau}\omega_1 G & 0 & 0 & \lambda \frac{1}{R_1^\tau}\omega_1 G' & 0 & \beta U_{L_1L_1} \end{array}\right).$$

When deriving the bordered Hessian, it was assumed that the cross derivatives of U are zero ($U_{CL} = 0$) and that human capital fully depreciates ($\delta_H = 1$). Dropping these simplifying assumptions does not change the following results.

Let D_k denote the kth leading principal minor. Then straightforward but tedious calculations yield

$$\begin{aligned} D_3 = & -(\omega_0 H_0)^2 U_{C_0C_0} - (-1)^2 U_{L_0L_0} > 0, \\ D_4 = & \lambda \frac{1}{R_1^\tau}\omega_1 L_1 G'' \big(-U_{L_0L_0} - U_{C_0C_0}(\omega_0 H_0)^2\big) < 0, \\ D_5 = & -\lambda \frac{1}{R_1^\tau}\omega_1 L_1 G'' \bigg(U_{C_0C_0} U_{L_0L_0} \left(-\frac{1}{R_1^\tau}\right)^2 + U_{L_0L_0}\beta U_{C_0C_0}(-1)^2 \\ & + U_{C_0C_0}\beta U_{C_1C_1}(\omega_0 H_0)^2 \bigg) > 0, \\ D_6 = & U_{C_0C_0}\beta U_{C_1C_1} U_{L_0L_0} \left(\frac{1}{R_1^\tau}\omega_1 G\right)^2 \left(-\lambda \frac{1}{R_1^\tau}\omega_1 L_1 G''\right) \\ & - \left(U_{L_0L_0}\beta U_{C_1C_1}(-1)^2 + U_{C_0C_0} U_{L_0L_0} \left(-\frac{1}{R_1^\tau}\right)^2 + U_{C_0C_0}\beta U_{C_1C_1}(\omega_0 H_0)^2 \right) \end{aligned}$$

$$\times \lambda \frac{1}{R_1^{\tau}} \omega_1 G'' U_{L_1} \frac{1}{1-\gamma} \big((1-\gamma) v_1 - \gamma \big) \overset{!}{<} 0$$

with $v_1 = L_1 U_{L_1 L_1} / U_{L_1}$ and $\gamma = \dfrac{G'}{G} E$.

The sign of D_6 must be negative. The minuend is negative. The first factor of the subtrahend is positive. Hence, the second factor must be positive:

$$(1-\gamma) v_1 > \gamma \Leftrightarrow \gamma < \frac{v_1}{1+v_1}.$$

The requirement is that the concavity of the utility function, captured by v_1, has to be sufficiently large to compensate for the lack of concavity of the law of motion (2) for human capital, measured by γ.