Int Tax Public Finance
DOI 10.1007/s10797-015-9357-0

Second-best income taxation and education policy with endogenous human capital and borrowing constraints

Bas Jacobs[1] · **Hongyan Yang**[2]

Abstract We formulate a two-period life cycle model of saving, labor supply, and human capital investment when individuals differ in their ability and initial wealth. Borrowing constraints result in sub-optimal choices for consumption and investments in human capital. We analyze optimal linear income taxes and education subsidies. The optimal income tax is shown to be positive—even in the absence of any redistributional concerns. A redistributive income tax relaxes borrowing constraints by redistributing resources from the unconstrained to the borrowing constrained stages of the life cycle. The income tax thus alleviates preexisting non-tax distortions in the capital market. Human capital is subsidized on a net basis in the absence of redistributional concerns. Education subsidies help to relax credit constraints and to reduce distortions from explicit and implicit taxes on human capital formation. When redistributional concerns are present, education is subsidized more if this helps to alleviate distortions on labor supply, but is subsidized less if education subsidies have a very regressive incidence. Simulations demonstrate that optimal income taxes are substantially higher when credit constraints are present. Education is generally subsidized on a net basis, and the more so if credit constraints are more severe.

Electronic supplementary material The online version of this article (doi:10.1007/s10797-015-9357-0) contains supplementary material, which is available to authorized users.

✉ Hongyan Yang
h.yang@hsu-hh.de

Bas Jacobs
bjacobs@ese.eur.nl
http://people.few.eur.nl/bjacobs

1 Erasmus School of Economics, Tinbergen Institute, and CESifo, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands

2 Helmut-Schmidt Universität Hamburg, Holstenhofweg 85, 22043 Hamburg, Germany

Published online: 12 May 2015

Keywords Labor taxation · Human capital investment · Credit constraints

JEL Classification H2 · I2 · J2

1 Introduction

This paper examines optimal income taxation, education subsidies, and human capital formation in an economy where individuals are subject to binding credit constraints. Empirical evidence for credit constraints is presented in two strands of the literature. First, (poor) individuals can experience difficulties financing their higher education, as shown in Kane (1996), Keane and Wolpin (2001), Plug and Vijverberg (2005), Belley and Lochner (2007), Stinebrickner and Stinebrickner (2008), and Lochner and Monge-Naranjo (2011).[1] Second, ample empirical evidence for binding borrowing constraints is found when empirically testing the life cycle hypothesis in consumption. See Attanasio and Weber (2010) for an excellent overview of this literature. Binding credit constraints could contribute to persistence in income mobility, result in larger inequality, strengthen segregation of neighborhoods, and decrease economic growth (Loury 1981; Galor and Zeira 1993; Durlauf 1996; Benabou 1996a, b; De Gregorio 1996; Mookherjee and Ray 2003; Galor and Moav 2004).

The purpose of this paper is to analyze optimal redistributive tax policies and optimal education policies when individuals cannot borrow sufficient funds to smooth consumption and to finance human capital investments. To that end, we develop a two-period life cycle model, where individuals make educational investments in the first period and they work in the second period of their life cycle. Exogenous constraints restrict the amount of borrowing that can be made by individuals in the first period of their life cycle, since poor individuals (or their parents) cannot collateralize human capital to finance investments in education (of their children). In a non-slave state, legal restrictions prevent individuals engaging in a contract that employs future income as collateral. However, the government can circumvent this constraint as it has a claim on all acquired human capital through the tax system (see also Stiglitz 1994; Jacobs and van Wijnbergen 2007). In addition, we allow the government to provide subsidies on education as education subsidies reduce the need to borrow and thereby indirectly alleviate credit constraints. Moreover, education subsidies can counter the distortions of the redistributive tax system on human capital investments, as in Bovenberg and Jacobs (2005). Indeed, our model extends their analysis of optimal income taxes and education subsidies by allowing for credit constraints.

The government sets linear income taxes and education subsidies so as to maximize social welfare. Ability and initial wealth are assumed to be non-verifiable. Hence, individualized lump-sum taxes are ruled out. Labor earnings and investments in human capital are verifiable so that income can be taxed and education can be subsidized. The government cannot perfectly eliminate all credit constraints or redistribute income

[1] Consistent with the presence of credit constraints, Kane (1995) and van der Klaauw (2002) identify large impacts of financial aid on college enrollment. Acemoglu and Pischke (2001) show that family income plays an important role in determining educational attainment. Caneiro and Heckman (2002) point out that credit constraints are relevant for about 8 % of the youth in the USA.

without causing distortions in labor supply and human capital investment. We demonstrate that the optimal income tax is progressive even in representative-agent settings where distributional concerns are absent. That is, we provide a case for distortionary taxation on efficiency grounds only. The intuition is that, as long as incomes are increasing over the life cycle, progressive tax systems redistribute resources from later and unconstrained stages to earlier and constrained stages in the life cycle. Hence, not only consumption is smoothed better, but also investments in human capital increase. The labor tax trades off the welfare gains of alleviating credit constraints against the tax distortions in labor supply and human capital formation. The optimal income tax rate is determined by the extent to which individuals are credit-constrained and the tax elasticities of labor supply and educational investment.

Moreover, we derive that education is subsidized on a net basis. Education subsidies are a useful complement to progressive income taxes to redistribute resources to the credit-constrained phase of the life cycle. However, since education is optimally subsidized on a net basis, education subsidies tend to create overinvestment in human capital, unlike lump-sum transfers. Therefore, the government uses both progressive income taxes and education subsidies to alleviate credit constraints. We demonstrate that education subsidies offset the explicit taxes on human capital formation by making costs of education effectively tax deductible. In addition, education subsidies offset the implicit taxes on human capital formation that originate from capital-market imperfections. Consequently, capital-market imperfections push optimal education subsidies beyond the point to make all investments tax deductible.

In a generalization of the model with heterogenous individuals, we show that the results derived under homogenous individuals carry over. With credit constraints, the trade-off between equity and efficiency is less severe, since redistribution now generates not only equity gains, but also efficiency gains. Hence, when distributional concerns are allowed for, the case for progressive income taxation is strengthened further. Moreover, like in the case with homogeneous agents, education is optimally subsidized to directly alleviate credit constraints and to reduce explicit and implicit taxes on human capital formation. However, education subsidies might be higher or lower in comparison with the representative-agent case for two other reasons. First, optimal education subsidies tend to be lower, because education subsidies are regressive. Second, education and labor supply are complementary in generating labor income. Consequently, education subsidies yield an efficiency gain by reducing some of the tax-induced distortions on labor supply. Whether education subsidies are higher or lower is not clear, since both effects work in opposite directions.[2] Only when capital markets are perfect, both effects exactly cancel out, and education is subsidized only to offset the distortions of the income tax on human capital formation. This result confirms Bovenberg and Jacobs (2005).

We simulate optimal taxes using an empirically plausible calibration of our model. Our simulations demonstrate that optimal taxes are more than 50 % higher when credit constraints are present compared to the optimal tax rates when credit constraints are absent. Moreover, the optimal tax rate in the absence of any redistributional concerns,

[2] Jacobs and Bovenberg (2011) show that similar mechanisms arise when the earnings function is not weakly separable as in Bovenberg and Jacobs (2005).

i.e., the efficient optimal tax rate, is as high as 43.6 %, indicating an important role of credit constraints for optimal taxation. Education is subsidized on a net basis in all simulations. Hence, there is overinvestment in human capital compared to the first-best rule. Education subsidies remove all distortions from explicit taxes and implicit taxes (due to capital-market imperfections) on skill formation. Moreover, we show that education subsidies are used primarily for efficiency reasons as the gains of lower labor-supply distortions largely cancel against the distributional costs in terms of a regressive incidence. Our simulation results are very robust to alternative specifications of the model.

Our paper relates in a number of ways to the existing literature. First, this paper analyzes credit constraints in optimal tax models with human capital investments. Almost the entire literature on human capital and optimal taxation assumes perfect capital markets or static human capital models in which capital markets play no substantive role, see, for example, Ulph (1977), Hare and Ulph (1979), Tuomala (1990), Eaton and Rosen (1980), Hamilton (1987), Nielsen and Sørensen (1997), Jones et al. (1993), Judd (1999), Anderberg and Andersson (2003), Jacobs (2005, 2012), Bovenberg and Jacobs (2005), Maldonado (2008), Boháček and Kapička (2008), Anderberg (2009), Jacobs and Bovenberg (2010, 2011), Schindler (2011), Jacobs et al. (2012), Findeisen and Sachs (2013), and Stantcheva (2014).

Second, we generalize the analyses in Bovenberg and Jacobs (2005), Maldonado (2008), Anderberg (2009), Boháček and Kapička (2008), Jacobs and Bovenberg (2010, 2011), Schindler (2011), Jacobs et al. (2012), Schindler and Yang (2015), Stantcheva (2014), among others, to analyze optimal redistributive income taxation joint with optimal education policies by allowing for capital-market imperfections. Besides countering explicit tax distortions, we show that education subsidies are also desirable to alleviate credit constraints and to counter implicit taxes on human capital due to capital-market failures. Moreover, higher education subsidies could be useful to alleviate labor-supply distortions, and lower education subsidies could be optimal to redistribute more income. It is not theoretically clear which of these two latter effects dominates in the presence of imperfect capital markets. However, our simulations demonstrate that these two effects are largely offsetting. A quantitative model with endogenous human capital and credit constraints is analyzed by Krueger and Ludwig (2013). Our theoretical analysis is in line with their simulation results, albeit in a simpler economic environment than theirs.

Third, our paper is related to Hubbard and Judd (1986), who simulate a life cycle model and demonstrate that progressive income taxation is welfare improving compared to proportional income taxation when credit constraints are binding. The intuition is the same as ours: the progressive income tax redistributes resources over the life cycle and allows for better consumption smoothing. Our paper, however, provides a formal proof for their finding as a special case of our model in which educational investment is kept exogenous.[3]

[3] Hoff and Lyon (1995) also show that redistributive income taxation improves welfare by mitigating adverse selection in the capital market. Taxing labor income progressively, and rebating the tax revenue through lump-sum transfers, increases collateralizable wealth. Progressive taxes thereby moderate

Fourth, earlier work by Loury (1981), Glomm and Ravikumar (1992), Benabou (1996a, b), and Fernandez and Rogerson (1996, 1998) demonstrates that when credit constraints are binding public provision of education or equalizing expenditure on education among communities can increase income equality, reduce segregation, promote income mobility, and boost economic growth. Tobin (1982) also points out that government policy should help credit-constrained individuals to move resources from the future to the present. However, this literature has not yet conducted an analysis of *optimal* redistributive policies when individuals face binding credit constraints.

Fifth, our paper also contributes to an extensive literature, which emphasizes the potentially efficiency-enhancing effects of distortionary taxes in second-best settings. See also van der Ploeg (2006a) for an overview. We show that the introduction of a distortionary tax instrument can reduce preexisting non-tax distortions in the economy arising from capital-market failures. For example, Akerlof (1976) shows that the introduction of a distortionary income tax helps to tame the 'rat race' and reduce individuals' excessive incentives to work. Related is Layard (1980, 2005) who argues that progressive taxation is welfare improving because individuals are involved in status races ('keeping up with the Joneses') and exhibit habit persistence, both giving excessive incentives to work. Labor-market imperfections arising from trade unions, efficiency wages, and search frictions also provide second-best arguments for progressive taxes (see Koskela and Vilmunen 1996; Pissarides 1998; Sørensen 1999; Boone and Bovenberg 2002; van der Ploeg 2006b; Bovenberg 2006). Unions set wages above market-clearing levels when unemployment benefits improve the outside options of workers. Also, firms pay too high efficiency wages in order to recruit, to retain, and to motivate workers when workers face attractive outside options. Progressive taxes punish both unions and firms to bid up wages, so that wages are moderated, and unemployment decreases.[4] Progressive taxation could also correct search frictions in labor markets. Progressive taxation lowers the wage demands by workers, which increases vacancies and expands employment. This is optimal if workers have too much bargaining power compared to firms, i.e., when the Hosios (1990) condition is not met. In the presence of missing insurance markets, progressive taxation redistributes income across different states of natures and improves efficiency by partially replacing the missing insurance market (Eaton and Rosen 1980; Varian 1980; Jacobs et al. 2012).

The remainder of this paper is organized as follows. Section 2 presents our life cycle model with imperfect credit markets and human capital investment. Optimal tax and education policies are analyzed in Sect. 3 in an economy with a representative individual, which focuses on optimal efficient taxation and education subsidies to relax borrowing constraints. In Sect. 4, we extend the model to a setting with heterogeneous individuals. Section 5 provides numerical simulations, and the last section concludes. The Appendix contains the proofs of all propositions. An online Appendix contains some derivations of intermediate results.

Footnote 3 continued
inefficient overinvestment in education. Our model, in contrast, emphasizes underinvestment in human capital resulting from binding credit constraints.

4 van Ewijk and Tang (2007) show that education subsidies are optimal in order to offset the disincentives on human capital investments when the government uses progressive taxes to lower union's wage demands.

2 Model

The economy is populated by a continuum of individuals living for two periods. The mass of all individuals is normalized to one. Individuals differ in their ability n and initial wealth ω. Ability and wealth have a cumulative joint distribution $F(n, \omega)$, which has supports $[\underline{n}, \infty)$ and $[\underline{\omega}, \infty)$. We use a subscript to denote the type of individual by its ability and initial wealth, while a superscript is used to label the period in the life cycle.

We consider a two-period life cycle model of investment in human capital, labor supply, saving, and borrowing constraints. In the first period, the individual does not work, but invests in education and consumes. In the second period, the individual supplies labor and consumes all its wealth.

Individuals derive utility from consumption in both periods, $c^1_{n\omega}$ and $c^2_{n\omega}$, and disutility from labor $l_{n\omega}$ in the second period. The utility function is assumed to be separable in consumption and labor:

$$U_{n\omega} \equiv u(c^1_{n\omega}, c^2_{n\omega}) - v(l_{n\omega}), \quad u_1, u_2, v' > 0, \quad u_{11}, u_{22}, -v'' < 0, \quad u_{12} \geq 0. \quad (1)$$

The sub-utility function $u(\cdot)$ is increasing and concave in both arguments. Moreover, we assume that sub-utility $u(\cdot)$ is homothetic. The subscripts refer to the derivatives with respect to the first and the second argument of the utility function, respectively. The disutility of labor $v(\cdot)$ is increasing and convex in $l_{n\omega}$.

Individuals invest $e_{n\omega}$ in human capital. We normalize the unit cost of consumption goods and educational investment to one. Investment in education is subsidized at rate s. Hence, $(1 - s)$ is the net unit cost of resources invested in human capital. Without loss of generality, we do not distinguish time and resource costs of education.[5] Second-period labor supply is denoted by $l_{n\omega}$. Gross labor income $z_{n\omega}$ depends on educational investment $e_{n\omega}$, labor supply $l_{n\omega}$, and ability n:

$$z_{n\omega} \equiv nl_{n\omega}\phi(e_{n\omega}) = nl_{n\omega}e^{\beta}_{n\omega}, \quad 0 < \beta < 1, \quad (2)$$

where $\phi(e_{n\omega})$ is the production function for human capital, which features positive, but diminishing marginal returns to human capital investment. Like Bovenberg and Jacobs (2005), we assume that the production function for human capital has a constant elasticity β. This assumption is made to ensure that optimal net taxes on education are zero in the absence of capital-market imperfections. All our results under capital-market imperfections can be generalized, however, to a general production function for human capital $\phi(e_{n\omega})$, with $\phi' > 0$ and $\phi'' < 0$.

Besides educational investment, the individual decides on its consumption in the first period $c^1_{n\omega}$ and saving $a_{n\omega}$. Consequently, the first-period budget constraint is:

$$a_{n\omega} = -(1 - s)e_{n\omega} + \omega + g - c^1_{n\omega}, \quad (3)$$

[5] As long as both investments are verifiable to the government and can be subsidized (or can be made tax deductible), this distinction would be immaterial for our results. See also Bovenberg and Jacobs (2005).

where g is the time-invariant lump-sum transfer. The (exogenous) interest rate equals r and is the same for saving and borrowing. Individuals are only allowed to borrow a maximum of a_o at the private capital market:

$$a_{n\omega} + a_o \geq 0. \tag{4}$$

This assumption reflects the fact that individuals have limited access to loans to finance consumption and educational investments. a_0 may also include an exogenously given level of government loans for the financing of education. Study loans are assumed to be insufficient to remove all capital-market imperfections, which is in accordance with the empirical evidence cited in the introduction. Since human capital is bad collateral, there exists a moral hazard problem in the repayment of study loans. Borrowing limits are then optimal to strengthen the incentives for loan repayment, but come at a cost of underinvestment in human capital. See also Andolfatto and Gervais (2006) and Lochner and Monge-Naranjo (2011).[6] Second-period consumption equals after-tax labor income, capital income, and the lump-sum transfer:

$$c_{n\omega}^2 = (1-t)nl_{n\omega}\phi(e_{n\omega}) + (1+r)a_{n\omega} + g, \tag{5}$$

where t denotes the labor tax rate.

The individual characteristics (n and ω), labor supply ($l_{n\omega}$), and saving decisions ($a_{n\omega}$) are assumed to be private information. In line with Mirrlees (1971), labor income $z_{n\omega}$ is verifiable to the government. Moreover, we follow Bovenberg and Jacobs (2005) by assuming that investment in education e_n is verifiable and can, therefore, be subsidized. We rule out individualized lump-sum taxes. The government has to rely on distortionary labor taxes to redistribute income. With a flat tax rate and positive non-individualized lump-sum transfers, the income tax is progressive. To simplify the exposition, we focus on the case of identical transfers in both periods, thereby ruling out age-dependent transfers. We demonstrate later that allowing for age-specific lump-sum transfers would only reinforce our findings.

The non-verifiability of ω implies that the government cannot levy taxes on initial wealth. This corresponds to the empirical observation that in most countries wealth is only lightly taxed, and in many countries, it is not taxed at all, see Eurostat, (2014).[7] Similarly, since saving is assumed to be non-verifiable, we rule out taxes on saving. As we will focus mainly on credit-constrained individuals, taxes on saving would not yield any revenues when savings are zero. Moreover, taxes on capital income do not have obvious value added over education subsidies. Both instruments boost human capital investment, but education subsidies avoid intertemporal distortions. Since all educational investment can be subsidized, optimal capital-income taxes are zero in the absence of capital-market imperfections due to separability and homotheticity in

[6] An extension of our model with endogenous borrowing limits under moral hazard is left for future research.

[7] Moreover, if we would assume that wealth is verifiable and allow for a wealth tax, it would be optimal to levy a confiscatory wealth tax, since initial wealth is exogenous. Our model would then become observationally equivalent to a model in which there is no heterogeneity in initial wealth. This is clearly counterfactual and removes the economically interesting correlation between ability and initial wealth.

preferences (Jacobs and Bovenberg 2010). Indeed, only known arguments for positive capital taxes in the presence of capital-market failures would be applicable, but these are unrelated to human capital formation.[8] Allowing for taxes on saving would thus not generate important additional insights.

We restrict our analysis to linear instruments, since our model features two-dimensional heterogeneity in both ability and initial wealth. Determining optimal non-linear policies in models with two-dimensional private information is generally not feasible.[9] By constraining the policy instruments to be linear, we are able to analyze a rich economic environment with two-dimensional heterogeneity.[10] The informational requirements for using flat-rate taxes and education subsidies are that the government only needs to verify aggregate labor income and aggregate investment in human capital.

The individual chooses educational investment $e_{n\omega}$, saving $a_{n\omega}$ and labor supply $l_{n\omega}$ to maximize utility (1) subject to the budget constraints (3), (5), and the credit constraint (4). After substituting the budget constraints (3) and (5) in the utility function (1), we can formulate the following Lagrangian $\mathcal{L}$ for the individual's maximization problem:

$$\max_{\{a_{n\omega},e_{n\omega},l_{n\omega}\}} \mathcal{L}_{n\omega} \equiv u\big(-(1-s)e_{n\omega}+\omega+g-a_{n\omega};\,(1-t)nl_{n\omega}\phi(e_{n\omega}) \\ +(1+r)a_{n\omega}+g\big)-v(l_{n\omega})+\varphi_{n\omega}(a_{n\omega}+a_o), \tag{6}$$

where $\varphi_{n\omega}$ is the Kuhn–Tucker multiplier on the credit constraint (4). The multiplier $\varphi_{n\omega}$ measures the marginal increase in individual utility if the individuals' borrowing limit a_o increases with one unit. The first-order conditions for utility maximization are given by

$$\frac{\partial\mathcal{L}_{n\omega}}{\partial a_{n\omega}} = -u_1(c^1_{n\omega}, c^2_{n\omega})+(1+r)u_2(c^1_{n\omega}, c^2_{n\omega})+\varphi_{n\omega}=0, \tag{7}$$

$$\varphi_{n\omega}\geq 0, \quad \varphi_{n\omega}=0 \text{ if } a_{n\omega}+a_o>0, \tag{8}$$

[8] Hubbard and Judd (1986) and İmrohoroğlu (1998) show that capital taxation improves welfare with binding credit constraints because capital taxation results in redistribution from unconstrained individuals (who do save) to credit-constrained individuals (who do not save). Consequently, credit constraints are alleviated, but this comes at a price of distorting the saving decisions of the non-constrained individuals. Aiyagari (1995) and Chamley (2001) show that capital taxation raises efficiency because borrowing constraints may generate "excessive saving" due to precautionary motives.

[9] From the mechanism-design literature, it is well known that optimal tax problems with two-dimensional asymmetric information are very complicated as one needs to take into account complex mimicking behaviors, i.e., double deviations, in two-dimensional type space. Recent examples solve such models by essentially eliminating the interactions between the two sources of heterogeneity in incentive-compatibility constraints. See, for example, Kleven et al. (2009) and Jacquet et al. (2013) who analyze random-participation models where one type of heterogeneity (participation costs) does not interfere with the incentive constraints in the dimension of the other source of heterogeneity (ability). Our model does not abide the formal structure of a random-participation model, and the analysis of non-linear instruments would be analytically non-tractable.

[10] Similarly, Piketty and Saez (2013) recently analyzed optimal piece-wise linear wealth and income taxes in models with two-dimensional heterogeneity in both earnings ability and bequests.

$$\frac{\partial \mathcal{L}_{n\omega}}{\partial e_{n\omega}} = -(1-s)u_1(c^1_{n\omega}, c^2_{n\omega}) + u_2(c^1_{n\omega}, c^2_{n\omega})(1-t)nl_{n\omega}\phi'(e_{n\omega}) = 0, \quad (9)$$

$$\frac{\partial \mathcal{L}_{n\omega}}{\partial l_{n\omega}} = u_2(c^1_{n\omega}, c^2_{n\omega})(1-t)n\phi(e_{n\omega}) - v'(l_{n\omega}) = 0. \quad (10)$$

If individuals are not credit-constrained ($\varphi_{n\omega} = 0$), the consumption and educational choices of the household can be summarized as

$$\frac{u_1(c^1_{n\omega}, c^2_{n\omega})}{u_2(c^1_{n\omega}, c^2_{n\omega})} = \frac{(1-t)nl_{n\omega}\phi'(e_{n\omega})}{(1-s)} = 1+r. \quad (11)$$

Intertemporal consumption choices are not distorted since the marginal rate of intertemporal substitution in consumption equals one plus the interest rate, which is the marginal rate of intertemporal transformation. The optimality condition for investment in education equates the marginal costs of investing one unit of resources in education $(1+r)$ with the marginal net benefits of one unit of resources invested in education $((1-t)nl_{n\omega}\phi'(e_{n\omega})/(1-s))$. Note that the marginal benefits of education increase if individuals supply more labor. Hence, labor and education are complements in generating gross income. If the marginal income tax rate t is larger (smaller) than the subsidy s on education, the tax system distorts educational investments downwards (upwards), since the future benefits of education are taxed at higher (lower) rates than current costs are subsidized.

For credit-constrained individuals ($\varphi_{n\omega} > 0$), we have $a_{n\omega} = -a_o$ and we obtain

$$\frac{u_1\left(c^1_{n\omega}, c^2_{n\omega}\right)}{u_2\left(c^1_{n\omega}, c^2_{n\omega}\right)} = \frac{(1-t)nl_{n\omega}\phi'(e_{n\omega})}{(1-s)} > 1+r. \quad (12)$$

The credit constraint creates a wedge in intertemporal consumption choices, i.e., a difference between marginal rate of intertemporal transformation $(1+r)$ and marginal rate of intertemporal substitution $\frac{u_1(\cdot)}{u_2(\cdot)}$, implying that individuals would like to transfer more consumption from the second to the first period if they could. Thus, a binding credit constraint renders income in the first period relatively more valuable to the individual than in the second period. Investment in education of credit-constrained individuals is distorted by the borrowing constraint, since the marginal returns to investment in human capital $((1-t)nl_{n\omega}\phi'(e_{n\omega})/(1-s))$ are larger than the marginal returns to financial saving $(1+r)$.

We can define the implicit tax $\pi_{n\omega}$ on human capital investment arising from the credit constraint as:

$$\pi_{n\omega} \equiv 1 - (1+r)\frac{u_2(\cdot)}{u_1(\cdot)}. \quad (13)$$

$\pi_{n\omega}$ measures to which extent the intertemporal consumption choices are distorted. An intertemporal consumption wedge implies that $\pi_{n\omega} > 0$, and $\frac{u_1(\cdot)}{u_2(\cdot)} > 1+r$. If the credit constraint is slack, there is no distortion caused by imperfect capital markets, i.e., $\pi_{n\omega} = 0$, and the standard Euler equation for consumption applies.

Using the definition of $\pi_{n\omega}$, the first-order condition for educational investment can be rewritten as

$$\frac{(1-\pi_{n\omega})(1-t)}{(1-s)} n l_{n\omega} \phi'(e_{n\omega}) = 1 + r. \tag{14}$$

From this equation, we can see that the binding credit constraint acts as an additional, implicit tax on the return from human capital investment. The value of $\pi_{n\omega}$ is different for individuals differing in both n and ω. In particular, $\pi_{n\omega}$ decreases with initial wealth until it becomes zero when individuals are no longer credit-constrained. It increases with ability n—for given levels of initial wealth ω, because more able individuals have a higher marginal return to education ($n l_{n\omega} \phi'(\cdot)$) and, consequently, would like to borrow more in order to finance larger investment in human capital.

Labor supply follows from the standard condition that equates the marginal rate of substitution between labor and consumption to the net marginal wage rate:

$$\frac{v'(l_{n\omega})}{u_2\left(c^1_{n\omega}, c^2_{n\omega}\right)} = (1-t) n \phi(e_{n\omega}). \tag{15}$$

Taxes distort labor supply by driving a wedge between the social rewards ($n\phi(e_{n\omega})$) and the private rewards ($(1-t)n\phi(e_{n\omega})$) of an additional hour work. Clearly, education and labor supply are complementary in generating labor income. Higher levels of education raise the marginal benefits of work as Eq. (15) demonstrates. Higher labor supply raises the marginal benefits from investing in human capital by increasing the utilization rate of human capital, see Eq. (14).

First-order conditions are necessary, but not sufficient due to the positive feedback between learning and labor supply. If we assume that the sub-utility function $u(\cdot)$ is homogeneous of degree one, the second-order condition requires $\beta(1+\varepsilon_{n\omega}) < 1$, where $\varepsilon_{n\omega} \equiv \left(\frac{v''(l_{n\omega})l_{n\omega}}{v'(l_{n\omega})}\right)^{-1}$ is the Frisch elasticity of labor supply.[11] A sufficiently low elasticity of labor supply $\varepsilon_{n\omega}$ and a sufficiently low elasticity of the human capital production function β ensure that the feedback between labor supply and education dampens out and interior solutions are obtained. We assume in the remainder that the second-order conditions are always respected.

The first-order conditions (14) and (15), the household budget constraints (3) and (5), and the credit constraint (4) jointly determine optimal investment in education, labor supply, saving and consumption choices as functions of the policy parameters g, t, and s, ability n, and initial wealth ω. By indicating the optimized values with an asterisk, we can write the indirect utility function as:

$$V_{n\omega}(g, t, s) \equiv u(c^{1*}_{n\omega}, c^{2*}_{n\omega}) - v(l^*_{n\omega}). \tag{16}$$

Applying Roy's identity yields the following derivatives with respect to the policy instruments: $\frac{\partial V_{n\omega}}{\partial g} = u_1(\cdot) + u_2(\cdot)$, $\frac{\partial V_{n\omega}}{\partial t} = -u_2(\cdot) n l_{n\omega} \phi(e_{n\omega})$, and $\frac{\partial V_{n\omega}}{\partial s} = u_1(\cdot) e_{n\omega}$.

For later reference, we also derive the Slutsky equations for education and labor supply in online Appendix B. With capital-market failures, deriving the compensated

[11] The derivation of the second-order conditions is provided in online Appendix A.

demand and supply functions is not trivial, because the exact timing of the compensation to keep utility fixed matters. If the credit constraint is slack, one unit of compensation given in first period is the same as the discounted value of one unit of compensation given in second period. However, if the credit constraint is binding, the value of one unit of compensation given in first period is higher than the discounted value of one unit of compensation in the second period. We derive the Slutsky equations where a uniform income compensation is given in *both* periods, that is, by a higher lump-sum transfer g:

$$\frac{\partial l_{n\omega}}{\partial t} = \frac{\partial l^c_{n\omega}}{\partial t} - \frac{u_2(\cdot)}{u_1(\cdot)+u_2(\cdot)} n l_{n\omega}\phi(e_{n\omega})\frac{\partial l_{n\omega}}{\partial g}, \tag{17}$$

$$\frac{\partial e_{n\omega}}{\partial t} = \frac{\partial e^c_{n\omega}}{\partial t} - \frac{u_2(\cdot)}{u_1(\cdot)+u_2(\cdot)} n l_{n\omega}\phi(e_{n\omega})\frac{\partial e_{n\omega}}{\partial g}, \tag{18}$$

$$\frac{\partial l_{n\omega}}{\partial s} = \frac{\partial l^c_{n\omega}}{\partial s} + \frac{u_1(\cdot)}{u_1(\cdot)+u_2(\cdot)} e_{n\omega}\frac{\partial l_{n\omega}}{\partial g}, \tag{19}$$

$$\frac{\partial e_{n\omega}}{\partial s} = \frac{\partial e^c_{n\omega}}{\partial s} + \frac{u_1(\cdot)}{u_1(\cdot)+u_2(\cdot)} e_{n\omega}\frac{\partial e_{n\omega}}{\partial g}. \tag{20}$$

where $l^c_{n\omega}$ denotes the compensated supply of labor, and $e^c_{n\omega}$ denotes the compensated demand for education.

3 Optimal taxation without redistribution

In this section, we discuss optimal income taxes and education subsidies when individuals are all identical and there are, consequently, no redistributional concerns. We therefore suppress the subscripts n and ω. Moreover, we assume that the initial wealth of the representative individual is not sufficient to finance the optimal level of education. Consequently, the credit constraint is binding and educational investment is inefficiently low. The case with a slack credit constraint is straightforward. Since all individual choices would be efficient, the government would have no need to intervene.

If age-specific lump-sum transfers (taxes) would be available in a setting with a representative individual, it would follow trivially that the credit constraint could be perfectly removed without any efficiency costs. In particular, a policy with age-specific transfers (taxes) can be viewed as a government loan where the government provides an amount of lump-sum income to each young individual and requires them to pay the loan back, including interest, with a lump-sum tax when they are old. Then, the government perfectly acts as a lender to replace the missing capital market without resorting to distortionary taxes on labor income. In heterogeneous-agent settings, which we will analyze in the next section, a first-best optimum, however, would require both age-specific *and* individualized lump-sum transfers (taxes), which are not feasible due to the informational constraints we have imposed on n and ω. The absence of individualized lump-sum transfers is therefore critical, not that taxes cannot be conditioned on age.

The tax system thus consists of a flat tax on labor income, a flat subsidy on education, and uniform lump-sum transfers in both periods. Without loss of generality, we assume that there are no exogenous government expenditures. We assume that the government is benevolent and has full commitment. That is, the government announces the tax and subsidy schedules before individuals make their decisions and fully commits to it.[12]

Tax revenue from labor taxation is used to finance education subsidies and lump-sum transfers. The government budget constraint is therefore given by:

$$tnl\phi(e) = (1+r)se + (2+r)g. \tag{21}$$

The tax payment in the second period should be equal to the value of education subsidies in the first period and transfers provided in both periods plus interest. Note that we express the government budget constraint in terms of second-period income. We assume that the government is not credit-constrained, as opposed to households. Intuitively, private markets will make government borrowing available, since the government can effectively collateralize human capital through the tax system. By the government's ability to tax income, the government can secure claims on future human capital returns (Jacobs and van Wijnbergen 2007). However, alleviating the credit constraint through the tax-transfer system is costly because labor supply is distorted, and the first-best allocation cannot be obtained.

The government maximizes the indirect utility of the representative agent $V(g, t, s)$ subject to its budget constraint (21). Optimal tax and education policy are given in the next proposition.

Proposition 1 (Optimal policy with a representative agent) *Optimal lump-sum transfers, income taxes, and education subsidies are determined by:*

$$b \equiv \frac{u_1(\cdot)+u_2(\cdot)}{\eta} + (tnl\phi'(e) - s(1+r))\frac{\partial e}{\partial g} + tn\phi(e)\frac{\partial l}{\partial g} = 2+r, \tag{22}$$

$$\frac{t}{1-t}(\varepsilon_{lt} + \beta\varepsilon_{et}) = \pi(1-\chi) + \frac{(1-\pi)s}{1-s}\beta\varepsilon_{et}, \tag{23}$$

$$\frac{(1-\pi)s}{1-s}\varepsilon_{es} = \pi\chi + \frac{t}{1-t}(\varepsilon_{ls}/\beta + \varepsilon_{es}), \tag{24}$$

where b is the social marginal value of income in monetary units, η is the shadow value of public resources, $\varepsilon_{lt} \equiv -\frac{\partial l^c}{\partial t}\frac{1-t}{l}$, $\varepsilon_{et} \equiv -\frac{\partial e^c}{\partial t}\frac{1-t}{e}$, $\varepsilon_{ls} \equiv \frac{\partial l^c}{\partial s}\frac{1-s}{l}$, and $\varepsilon_{es} \equiv \frac{\partial e^c}{\partial s}\frac{1-s}{e}$ denote the compensated elasticities of labor supply and educational investment with respect to the taxes and subsidies, and $1-\chi \equiv \frac{1+r}{2+r-\pi} = \left(1+\frac{u_2(\cdot)}{u_1(\cdot)}\right)^{-1}$ is a measure for the imperfection of an age-independent tax system.

Proof See Appendix A. □

[12] However, in view of the sunk character of the educational investment, the optimal policy is generally not time-consistent. Therefore, a benevolent government may want to renege on its announcements and re-optimize taxes after investments have been made, see also Pereira (2009).

3.1 Optimal transfers

Equation (22) states that the marginal social benefit b of providing one unit of income in both periods (including the indirect income effects on the tax bases) should be equal to the marginal resource cost of providing one unit of income in both periods (see also Atkinson and Stiglitz 1980). The optimal lump-sum transfer thus equalizes the social marginal value of resources in the public and the private sector.

3.2 Optimal income taxes

Equation (23) gives the first-order condition for the optimal income tax when transfers g are optimized, for a given level of the education subsidy s. The optimal tax trades off the marginal welfare gains of alleviating the credit constraint—captured by the first term on the right-hand side—against the marginal efficiency costs of doing so—captured by the term on the left-hand side and the second term on the right-hand side. In the optimum, the welfare gains of alleviating the credit constraint should be equal to their efficiency costs.

A redistributive income tax relaxes credit constraints, because it transfers income from non-constrained toward constrained phases in the life cycle. The more individuals are credit-constrained, as measured by a higher value of $(1-\chi)\pi$ on the right-hand side of (23), the larger is the welfare gain of a higher tax rate. $\chi \equiv \frac{1-\pi}{2+r-\pi}$ measures the inefficiency of an age-independent tax system compared to a tax system where the transfer is provided in the first period only. In the latter case, we would obtain $\chi = 0$. Intuitively, for given tax and subsidy rates (and, therefore, for a given level of efficiency costs) the resources available to be transferred to the first period are lower when the same amount has to be transferred to the second period as well. Hence, for one unit of tax revenue raised in second period by the labor tax, only $\frac{1+r}{2+r}$ can be transferred to the first period. Due to the 'leak' of the transfers to the second period, the credit constraint is alleviated to a lesser extent, and the optimal tax rate is lower as a result. The relative share of tax revenue that can be transferred to the first period increases if a higher interest rate (r higher) or more severe capital-market failures (π higher) make intertemporal transfers less costly to the government than to households. The reason is that government faces a lower relative price for first-period consumption, i.e., $1+r$, than households, i.e., $\frac{1+r}{1-\pi}$.[13]

The larger is deadweight loss of income taxation on the left-hand side of (23), the lower should be the optimal tax rate. Marginal deadweight loss consists of the marginal excess burden of taxes on labor income, i.e., $\frac{t}{1-t}\varepsilon_{lt}$, and the marginal excess burden of taxes on human capital investment, i.e., $\frac{t}{1-t}\beta\varepsilon_{et}$. The higher are the compensated tax elasticities of labor supply ε_{lt} and education $\beta\varepsilon_{et}$, the larger are tax distortions, and the lower should the optimal tax rate be.

[13] If we would allow for age-specific transfers, the optimal policy would be to provide transfers in the first period only, such that the inefficiency parameter χ is reduced to zero. This implies a larger efficiency gain from alleviating credit constraints through redistributive taxation, and our result of optimal progressive taxation would be strengthened further.

Education subsidies reduce the efficiency costs of the income tax, as indicated by the last term on the right-hand side of (23). The higher is the education subsidy, the more education subsidies reduce income tax distortions on education. We see that education subsidies are less effective to alleviate the explicit tax on learning if the implicit tax π on human capital due to capital-market imperfections is larger.

By eliminating the education subsidy from (23), using (24), we can derive that optimal income taxes are positive in the tax optimum:

$$\frac{t}{1-t} = \left(\frac{1 - \chi + \chi\beta\frac{\varepsilon_{et}}{\varepsilon_{es}}}{\varepsilon_{lt} - \frac{\varepsilon_{ls}}{\varepsilon_{es}}\varepsilon_{et}} \right) \pi > 0. \tag{25}$$

Income taxes are higher if credit constraints are more severe (i.e., a higher π). Hence, even in the absence of redistributional concerns, the optimal labor tax rate is positive. We thus provide a second-best argument for employing distortionary income taxation for efficiency reasons. The distortionary income tax helps to reduce a preexisting non-tax distortion in capital markets. See also the introduction for references to the literature on efficient income taxation in models with rat-races and habit persistence, distorted labor markets, and missing insurance markets.

The elasticity of the total tax base is given by in the denominator of (25). Optimal income taxes are lower when the tax elasticity of labor supply ε_{lt} increases. Education subsidies reduce the elasticity of taxable income by the term $\frac{\varepsilon_{ls}}{\varepsilon_{es}}\varepsilon_{et}$. This term captures how the optimal tax system exploits the complementarity between learning and working. The optimal income tax is higher if education subsidies are more useful to alleviate labor-market distortions. Consequently, income taxes are higher if labor supply responds elastically to education subsidies (ε_{ls} high) so that labor-supply distortions decrease a lot when education is subsidized more. Income taxes are lower if education responds elastically (ε_{es} high) so that human capital distortions increase a lot when subsidies are higher.

The optimal income tax is also higher when education subsidies are more effective in alleviating credit constraints, cf. $\chi\beta\frac{\varepsilon_{et}}{\varepsilon_{es}}$ in the numerator (see also our discussion in Sect. 3.3). In that case, education subsidies are optimally higher, which lowers the net tax on human capital. Consequently, net tax distortions in human capital formation are lower, and income taxes can be set higher, accordingly.

If investment in education and labor supply would both become perfectly inelastic ($\varepsilon_{lt} = \varepsilon_{et} = 0$), the labor tax would become completely non-distortionary and the first-best allocation would be obtained. The labor tax then becomes a second-period lump-sum tax, which differs from the first-period lump-sum transfer g. Consequently, the government can perfectly mimic a non-distortionary loan system with income taxes and transfers.

If there are no education subsidies (i.e., $s = 0$), the optimal income tax follows by substituting $s = 0$ in (23) (see also Jacobs and Yang 2013):

$$\frac{t}{1-t} = \frac{\pi(1-\chi)}{\varepsilon_{lt} + \beta\varepsilon_{et}}. \tag{26}$$

This case most clearly demonstrates that income taxes are positive for efficiency reasons alone. Our efficiency case for progressive income taxation does not rely exclusively on the endogeneity of human capital investments. Indeed, the optimal income tax would be progressive even when human capital would be exogenous (i.e., $\beta\varepsilon_{et} = 0$). Therefore, we formally prove the numerical findings by Hubbard and Judd (1986) that optimal income taxes are progressive when individuals face borrowing constraints.

Furthermore, it can be seen that our results also do not rely on the assumption of fully non-deductible costs of education, i.e., only resource costs. If education decisions are not distorted at all by the tax-subsidy system (i.e., $\frac{(1-\pi)s}{1-s} = \frac{t}{1-t}$), we have:

$$\frac{t}{1-t} = \frac{\pi(1-\chi)}{\varepsilon_{lt}}. \tag{27}$$

Hence, the case for progressive income taxes is not lost even if education subsidies remove all inefficiencies in human capital formation. Income taxes remain positive and would even be higher than in the case of non-deductible costs, since the elasticity of the tax base would be smaller for a given distortion of the credit constraint $\pi(1-\chi)$.

3.3 Optimal education subsidies

Equation (24) gives the first-order condition for the optimal education subsidy when transfers g are optimized, for a given level of the income tax t. The left-hand side of (24) represents the marginal deadweight loss of distorting human capital investment. The effective subsidy equals $\frac{(1-\pi)s}{1-s}$. Implicit taxes on human capital π thus reduce the impact of education subsidies on skill formation. Education subsidies are optimally higher if tighter credit constraints (larger π) reduce the effective subsidy rate more, ceteris paribus the right-hand side of (24). The optimal education subsidy is lower if human capital decisions are more elastic, i.e., ε_{es} is larger. In that case, subsidies on education create larger distortions in skill formation. The right-hand side of (24) gives the marginal benefits of education subsidies. We see that education subsidies have two benefits.

First, education subsidies help to alleviate the credit constraint as captured by $\pi\chi$. Education subsidies should be higher if credit constraints are more severe, i.e., π is higher. Since education subsidies are targeted exclusively at the first period of the life cycle, they are more useful to alleviate the liquidity constraint than lump-sum transfers are, but lump-sum transfers create fewer distortions than education subsidies. In particular, education subsidies are optimally set beyond levels to ensure full efficiency in skill formation, see below. Hence, there will be overinvestment in human capital, which can be avoided by using lump-sum transfers. The larger is χ, the more lump-sum transfers 'leak' to the second period of the life cycle, see our discussion above. Consequently, education subsidies are more desirable. If all lump-sum transfers could be perfectly targeted to the first period, $\chi = 0$, there would be no role for education subsidies at all to alleviate credit constraints. Intuitively, the government then prefers to use the least distortionary instrument to relax credit constraints.

Second, the second term in Eq. (24) shows that the subsidy alleviates tax distortions on labor supply and human capital formation as captured by $\frac{t}{1-t}(\varepsilon_{ls}/\beta + \varepsilon_{es})$. If education subsidies are effective to boost labor supply, ε_{ls}/β is large, and education subsidies are higher to reduce distortions on labor supply. Similarly, if education subsidies are effective to increase human capital investment, ε_{es} is large, and optimal subsidies should be higher to offset distortions on human capital investment.

In order to capture the full impact of the entire tax-subsidy system on human capital formation, we can define the net tax wedge on education as (cf. Bovenberg and Jacobs 2005):

$$\Delta \equiv \frac{t}{1-t} - \frac{(1-\pi)\,s}{1-s}. \tag{28}$$

Δ determines whether human capital investment is taxed ($\Delta > 0$) or subsidized ($\Delta < 0$) on a net basis. Higher income taxes t distort education downwards by depressing its returns. Higher education subsidies s distort education upwards by lowering its cost. Implicit taxes π raise the total tax wedge Δ. By substituting (23) and (24) in (28), we find that education is optimally subsidized on an net basis in the tax optimum:

$$\beta\Delta = -\left(\frac{\beta\chi + (1-\chi)\frac{\varepsilon_{ls}}{\varepsilon_{lt}}}{\varepsilon_{es} - \frac{\varepsilon_{et}}{\varepsilon_{lt}}\varepsilon_{ls}}\right)\pi < 0. \tag{29}$$

Since $\Delta < 0$, the optimal net tax wedge on education is negative, and education is optimally subsidized beyond the point to remove all tax distortions. The optimal net subsidy on learning Δ is larger if capital-market imperfections are more important (higher π). Consequently, some overinvestment in human capital is socially desirable—relative to the first-best decision rule. Education is not subsidized on a net basis when the capital market works perfectly ($\pi = 0$). Moreover, income taxes are zero then as well.

Net subsidies on education increase if they are more desirable to soften credit constraints (i.e., $\beta\chi$ is large) and decrease if education gets more severely distorted (i.e., ε_{es} is larger). However, net education subsidies increase if income taxes eliminate part of the overinvestment, since the total elasticity of the net subsidy base declines in $\frac{\varepsilon_{et}}{\varepsilon_{lt}}\varepsilon_{ls}$. This term captures, again, the complementarity between learning and working. Income taxation is more powerful to reduce overinvestment in human capital when education responds more elastically to taxation (i.e., ε_{et} is higher), and labor supply is not very elastic with respect to taxation (i.e., ε_{lt} is low).

The optimal net tax on education Δ decreases when income taxes are more effective to alleviate credit constraints, i.e., when the term in the numerator $(1-\chi)\frac{\varepsilon_{ls}}{\varepsilon_{lt}}$ is larger. As income taxes are then optimally set at higher rates, labor supply is more severely distorted. To alleviate these larger labor tax distortions, higher net subsidies on education are optimal.

4 Optimal taxation with redistribution

In this section, we allow the individuals to differ in their initial wealth ω and their innate ability n. By doing so, we introduce redistributional concerns in the optimal-tax problem. Since we assume that neither ω nor n are observable to the government, individualized lump-sum transfers that are conditioned on either ability or initial wealth are excluded. Consequently, the government has to rely on distortionary labor income taxation to redistribute income between individuals. Moreover, we allow the government to use education subsidies to offset tax distortions and alleviate credit constraints. Like before, revenues from the labor tax are used to finance education subsidies and non-individualized lump-sum transfers in both periods.

The government maximizes a social welfare function, which is a sum of concave individual indirect utilities $V_{n\omega}(g, t, s)$:

$$\int_{\underline{n}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi(V_{n\omega}(g, t, s))\mathrm{d}F(n, \omega), \quad \Psi' > 0, \quad \Psi'' \leq 0. \tag{30}$$

The social welfare function is utilitarian if $\Psi' = 1$, and it is Rawlsian if it features $\Psi' = 0$ for all individuals, except for the individual with the lowest utility. The government budget constraint is given by:

$$\int_{\underline{n}}^{\infty} \int_{\underline{\omega}}^{\infty} [tnl_{n\omega}\phi(e_{n\omega}) - (1+r)se_{n\omega}]\,\mathrm{d}F(n, \omega) = (2+r)g. \tag{31}$$

In order to characterize the optimal policies, we follow Diamond (1975) by defining the net social marginal valuation of one unit of income in both periods, measured in monetary units, for individuals with ability n and initial wealth ω as:

$$b_{n\omega} \equiv \frac{\Psi'(V_{n\omega})(u_1(\cdot) + u_2(\cdot))}{\eta} + tn\phi(e_{n\omega})\frac{\partial l_{n\omega}}{\partial g} + \left(tnl_{n\omega}\phi'(e_{n\omega}) - s\,(1+r)\right)\frac{\partial e_{n\omega}}{\partial g}. \tag{32}$$

Furthermore, we define the distributional characteristics of labor income and education as (see also Atkinson and Stiglitz 1980):

$$\xi_z \equiv -\frac{\mathrm{cov}\left[z_{n\omega}, \frac{b_{n\omega}}{2+r}\right]}{\bar{z}\frac{\bar{b}}{2+r}} = \frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left(1 - \frac{b_{n\omega}}{2+r}\right) z_{n\omega}\mathrm{d}F(n, \omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} z_{n\omega}\mathrm{d}F(n, \omega)\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{b_{n\omega}}{2+r}\mathrm{d}F(n, \omega)} > 0. \tag{33}$$

$$\xi_e \equiv -\frac{\mathrm{cov}\left[e_{n\omega}, \frac{b_{n\omega}}{2+r}\right]}{\bar{e}\frac{\bar{b}}{2+r}} = \frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left(1 - \frac{b_{n\omega}}{2+r}\right) e_{n\omega}\mathrm{d}F(n, \omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} e_{n\omega}\mathrm{d}F(n, \omega)\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{b_{n\omega}}{2+r}\mathrm{d}F(n, \omega)} > 0. \tag{34}$$

where $\bar{z} \equiv \int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} z_{n\omega}\mathrm{d}F(n, \omega)$, $\bar{e} \equiv \int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} e_{n\omega}\mathrm{d}F(n, \omega)$, and $\bar{b} \equiv \int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} b_{n\omega}$ $\mathrm{d}F(n, \omega)$, denote average income, average education, and average social welfare weight of individuals, respectively. We rescaled the welfare weights $b_{n\omega}$ by $2 + r$. This ensures that the scaled welfare weights have a mean equal to one.

The distributional characteristic ξ_z measures the marginal increase in social welfare, expressed in monetary units as a fraction of taxed labor income, of a marginally higher tax on labor earnings. ξ_z is the (negative) normalized covariance between the scaled welfare weight, $b_{n\omega}/(2+r)$, and gross labor income, $z_{n\omega}$. A positive value of ξ_z implies that individuals with a higher gross labor income $z_{n\omega}$ have a lower social welfare weight $b_{n\omega}$. Consequently, taxing labor income gives redistributional benefits. Similarly, ξ_e captures the marginal increase in social welfare, expressed in monetary units as a fraction of investment in education, of a marginal redistribution of income by a higher net tax on education. We presume that the individuals with lower social welfare weights (e.g., due to a higher ability or more initial wealth) invest more in education. Hence, the distributional characteristic of education ξ_e is positive, implying that (net) taxes on education provide distributional benefits. Equivalently, subsidies on education thus result in redistributional losses by transferring resources to those individuals with low welfare weights. Social marginal welfare weights decline due to diminishing marginal utility of income at the private level and/or concavity of the social welfare function, cf. the definition of $b_{n\omega}$ in (32). The ξ-terms for labor income and education are, therefore, measures for the strength of redistributional concerns implied by the social welfare function (30). The ξ-terms are zero if the social welfare weights $b_{n\omega}$ for all individuals are equal, so that the government does not want to redistribute any income by taxing income or education. Similarly, when there is no inequality in labor earnings $z_{n\omega}$ (education $e_{n\omega}$), taxing income (education) does not provide any redistributional benefits.

The government maximizes social welfare (30) by optimally setting transfers g, income taxes t, and education subsidies s, subject to its budget constraint (31). The next proposition characterizes optimal tax and education policies with heterogeneous agents.

Proposition 2 (Optimal policy with heterogeneous agents) *Optimal lump-sum transfers, income taxes, and education subsidies are determined by:*

$$\overline{b} = 2 + r, \tag{35}$$

$$\frac{t}{1-t}(\overline{\varepsilon_{lt}} + \beta\overline{\varepsilon_{et}}) = \xi_z + \frac{\overline{\pi(1-\chi)b}}{\bar{b}} + \frac{s}{1-s}\beta\overline{(1-\pi)\varepsilon_{et}}, \tag{36}$$

$$\frac{s}{1-s}\overline{(1-\pi)\varepsilon_{es}} = -\overline{(1-\pi)}\xi_e + \frac{\overline{\pi\chi b}}{\bar{b}} + \frac{t}{1-t}(\overline{\varepsilon_{ls}}/\beta + \overline{\varepsilon_{es}}). \tag{37}$$

where the bars indicate income-weighted averages of all variables x, i.e., $\bar{x} \equiv \left[\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} x_{n\omega} z_{n\omega} \mathrm{d}F(n,\omega)\right]\left[\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} z_{n\omega}\mathrm{d}F(n,\omega)\right]^{-1}$*, except for $\bar{b}$ which is not weighted with income.*

Proof See Appendix A. □

Compared with the results in Proposition 1, there are two main differences in the optimal tax expressions with heterogeneous agents. First, and foremost, we see in Eq. (36) that income taxes have distributional benefits ξ_z. Of course, this does not come as a surprise; the larger are the distributional benefits of income taxation, the higher should the income tax rate be. Education subsidies then also rise, since they optimally reduce tax distortions in human capital formation and labor supply.

Similarly, in Eq. (37), we see that education subsidies now entail distributional losses as the term $-\overline{(1-\pi)\xi_e}$ demonstrates. The more regressive education subsidies are, the lower education subsidies should optimally be, ceteris paribus the benefits of education subsidies to alleviate credit constraints (i.e., $\overline{\pi\chi b}/\bar{b}$) and to eliminate tax distortions (i.e., $t/(1-t)(\overline{\varepsilon_{ls}}/\beta+\overline{\varepsilon_{es}})$). Moreover, when education subsidies are set at lower levels, income taxes need to decline as well, since these create larger distortions in labor supply and skill formation. Except for the distributional terms, all the terms in the optimal tax formulae (35), (36), and (37) are completely analogous to the ones we discussed before in the case with a representative agent. Hence, the interpretation of these will not be repeated again.

Our findings with heterogeneous agents demonstrate again that taxing labor income can improve efficiency by correcting under-investment in education and reducing the intertemporal distortions in consumption due to credit constraints. As a result, the trade-off between equity and efficiency is less severe. Indeed, even without any desire for income redistribution ($\xi_z = 0$), the labor tax rate optimally remains positive, implying a case for distortionary taxation only on grounds of efficiency only, as we have demonstrated above for homogenous individuals.

The second, relatively minor, difference compared to the representative-agent case is the income-weighting for all elasticities and the π- and χ-parameters. And, therefore, the optimal tax and subsidy rates are not the same as those in (23) and (24) even if there are no distributional concerns ($\xi_z = \xi_e = 0$). The tax distortions and credit constraints of high-income earners are relatively more important than the tax distortions and credit constraints of low-income earners, because their elasticities and credit constraints are weighted more heavily.

In general, we cannot analytically solve for the income-weighted average of net taxes on education $\bar{\Delta}$ due to the interactions between implicit taxes $\pi_{n\omega}$ and the behavioral elasticities $\varepsilon_{lt,n\omega}$, $\varepsilon_{ls,n\omega}$, $\varepsilon_{et,n\omega}$ and $\varepsilon_{es,n\omega}$. However, we gain more economic intuition into the main mechanisms that drive optimal net taxes on education if we make the assumption that the behavioral elasticities $\varepsilon_{lt,n\omega}$, $\varepsilon_{ls,n\omega}$, $\varepsilon_{et,n\omega}$ and $\varepsilon_{es,n\omega}$ are uncorrelated with implicit taxes $\pi_{n\omega}$. Then, the income-weighted average of net taxes on education $\bar{\Delta}$ follows from taking the income-weighted average of (28) and substituting (36) and (37) to find (see Appendix B):

$$\beta\bar{\Delta} = -\left(\frac{\frac{\overline{\chi\pi b}}{\bar{\pi}\bar{b}}\beta + \frac{\overline{(1-\chi)\pi b}}{\bar{\pi}\bar{b}}\frac{\overline{\varepsilon_{ls}}}{\overline{\varepsilon_{lt}}}}{\overline{\varepsilon_{es}} - \frac{\overline{\varepsilon_{et}}}{\overline{\varepsilon_{lt}}}\overline{\varepsilon_{ls}}}\right)\bar{\pi} - \left(\frac{\frac{\xi_z}{\overline{(1-\pi)\xi_e}} - \frac{\overline{\varepsilon_{lt}}}{\overline{\varepsilon_{ls}}/\beta}}{\overline{\varepsilon_{es}}\frac{\overline{\varepsilon_{lt}}}{\overline{\varepsilon_{ls}}} - \overline{\varepsilon_{et}}}\right)\overline{(1-\pi)\xi_e} \gtrless 0. \quad (38)$$

From expression (38) follows that the first term is the same as in formula (29) except for the income-weighting of all terms under heterogeneous agents. Hence, for this reason credit constraints call for a net subsidy on human capital formation, as discussed above. The main difference with the no-redistribution case is the presence of the second term. The term in the numerator (i.e., $\xi_z/(\overline{(1-\pi)\xi_e}) - \beta\overline{\varepsilon_{lt}}/\overline{\varepsilon_{ls}}$) shows that labor-supply distortions and redistributional concerns come into play as well in determining the optimal net tax on education.

First, redistributive concerns call for net taxes on human capital. Whether net taxes on education are employed depends on whether they redistribute relatively more (less) income in comparison with income taxes. If taxes on income redistribute more (less) income than net taxes on education, i.e., $\xi_z > (<) \xi_e \overline{(1-\pi)}$, human capital formation should be subsidized more (less) on a net basis—ceteris paribus.

Second, efficiency concerns may call for net subsidies on human capital. Typically, education boosts labor supply, since learning and working are complements in generating income. Therefore, education subsidies can alleviate some of the tax-induced distortions on labor supply. If net taxes on education distort labor supply relatively more (less) than income taxes do, i.e., $\overline{\varepsilon_{ls}}/\beta > (<) \overline{\varepsilon_{lt}}$, then education should be subsidized relatively more (less)—ceteris paribus. Taxes on income and education are equivalent only if $\xi_e \overline{(1-\pi)}/\xi_z = (\overline{\varepsilon_{ls}}/\beta)/\overline{\varepsilon_{lt}}$. In that case, the redistributional (dis)advantage of net taxes on education over income taxes is exactly offset by the larger (smaller) relative distortions of education taxes over income taxes in the labor market.[14]

When credit constraints are absent, the optimal income tax and education subsidies collapse to the standard expressions for optimal linear income taxation and education subsidies with endogenous human capital formation. In particular, the optimal income tax then equals $t/(1-t) = \xi_z/\overline{\varepsilon_{lt}}$, and education subsidies equal tax rates, i.e., $s = t$, so that education is not taxed on a net basis, i.e., $\bar{\Delta} = 0$. See Appendix C for the proof. These findings confirm Bovenberg and Jacobs (2005). Intuitively, education subsidies only play a role to offset the distortions of the income tax on skill formation. Due to the linearity between education and earnings, net taxes on education are distributionally equivalent to taxes on labor income (i.e., $\xi_z = \xi_e$), but, in addition, also distort human capital investment. For the same reason, a tax on education has the same labor-supply distortions as a tax on labor income (i.e., $\overline{\varepsilon_{ls}}/\beta = \overline{\varepsilon_{lt}}$), but in addition also distorts education. Education subsidies are, therefore, neither used to alleviate labor-supply distortions, nor to redistribute income. See also Jacobs and Bovenberg (2011).

When there are capital-market imperfections, the borrowing constraint breaks the linear relationship between earnings and education. Therefore, education needs to be subsidized at a lower (higher) rate than the income tax rate—ceteris paribus—if investment in human capital increases more (less) than proportionally in individuals' earnings, i.e., when $\xi_z < (>) \xi_e$. Similarly, it is no longer clear whether income taxes harm labor supply more than net education taxes, since $\overline{\varepsilon_{lt}}$ is different from $\overline{\varepsilon_{ls}}/\beta$. Online Appendix C formally derives both elasticities and which one is larger depends on the particular parameters of the model. Therefore, we cannot conclude in which direction optimal education subsidies should be adjusted to alleviate labor-market distortions.

5 Numerical examples

In this section, we provide some numerical simulations to demonstrate the effects of credit constraints on optimal income taxes and education subsidies. The simulations

[14] See also Jacobs and Bovenberg (2008) who show that very similar mechanisms are present under optimal linear taxes and education subsidies under perfect capital markets and general earnings functions.

require information on the joint distribution of abilities and initial wealth, the utility function, and the social welfare function.

Ability and initial wealth are assumed to be jointly log-normally distributed with correlation coefficient ϱ between log-abilities and log-wealth levels. Log-normal distributions are useful first approximations to real-world income and wealth distributions. Log-ability is normally distributed with mean μ_n and standard deviation σ_n. Similarly, log-wealth is normally distributed with mean μ_ω and standard deviation σ_ω. We construct a data set representing the deciles of each distribution to form 100 ability-wealth classes for the entire population. We assume that initial wealth and ability are positively correlated and set $\varrho = 0.3$ for the benchmark case. The correlation coefficient is chosen such that about 20 % of the population is credit-constrained in the baseline simulation, see below.

We assume that the utility function features a CES sub-utility function for consumption and an iso-elastic sub-utility function for labor:

$$U_{n\omega} \equiv \left[\alpha (c^1_{n\omega})^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(c^2_{n\omega})^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma\nu}{\sigma-1}} - \gamma \frac{l_{n\omega}^{1+1/\varepsilon}}{1+1/\varepsilon}, \tag{39}$$
$$\gamma, \varepsilon > 0, \quad 0 < \alpha < 1, \quad 0 < \nu \leq 1, \quad \sigma > 0.$$

σ is the constant elasticity of intertemporal substitution in consumption. Based on the empirical evidence provided in Guvenen (2006) and Attanasio and Weber (2010), we assume an intertemporal elasticity of substitution of $\sigma = 0.67$ for the benchmark case. We assume an annual interest rate of 3 % and that one period in our model corresponds to 25 years. Hence, the interest rate r in our model equals $(1.03)^{25} - 1 = 1.09$. Jointly with the intertemporal elasticity of substitution σ and the interest rate r, the parameter α determines the life-time consumption profile. We calibrate $\alpha = 0.42$ such that second-period consumption is twice as high as first-period consumption in the benchmark calibration.

The parameter ν captures income effects in labor supply and the education response to imperfect capital markets. Income effects in labor supply are absent for $\nu = 1$, since the sub-utility function for consumption is then linearly homogenous, so that the marginal utility of income becomes constant. The benchmark case assumes $\nu = 0.25$ to allow for income effects. The Frisch elasticity of labor supply is set at $\varepsilon = 0.25$, which we approximated by the estimated compensated labor-supply elasticities for men that lie between 0.2 and 0.3 on average.[15]

The production function of human capital is assumed to be Cobb-Douglas:

$$\phi(e_{n\omega}) = e_{n\omega}^{\beta}, \quad 0 < \beta < 1. \tag{40}$$

[15] Extensive surveys by Blundell and MaCurdy (1999) and Evers et al. (2008) document that the uncompensated labor-supply elasticity is on average between 0 and 0.1 for men and around 0.5 for women. The estimates for the income elasticity in Blundell and MaCurdy (1999) indicate a value of around 0.2 on average for both men and women.

The elasticity of the human capital production function is set at $\beta = 0.5$ for the benchmark case, which is the average of $\beta = 0.6$ used in Trostel (1993) and $\beta = 0.4$ used in Jacobs (2005).

The social welfare function is a Bergson-Samuelson function with a constant elasticity of inequality aversion ζ, see e.g., Atkinson and Stiglitz (1980):

$$\Psi(V_{n\omega}) = \begin{cases} \frac{V_{n\omega}^{1-\zeta}-1}{1-\zeta}, & \zeta \neq 1, \\ \ln V_{n\omega}, & \zeta = 1. \end{cases} \tag{41}$$

For $\zeta = 0$, the social welfare function is utilitarian. For $\zeta = \infty$, the welfare function converges to the Rawlsian maxi-min case. For our benchmark simulation, we calibrate ζ to obtain optimal tax rates that are in line with commonly observed real-world values. Incidentally, it turns out that a utilitarian objective (i.e., $\zeta = 0$) gives reasonable optimal tax rates.

To calibrate our model, we assume an observed effective marginal tax rate of 48.8 %.[16] We assume an effective education subsidy of 56.6 % on total investment in human capital.[17] We assume a higher standard deviation of initial wealth than that of ability given that wealth distributions are typically more unequal than income distribution. Moreover, we assume that the borrowing limit is zero, i.e., $a_o = 0$. The remaining parameters—the distribution parameters μ_n, μ_ω, σ_n, and σ_ω, and dis-utility of labor γ—are jointly calibrated to meet 3 conditions that we impose on the model. (i) Mean labor supply equals 0.5. This value is taken from Jacobs (2005), Stern (1976), and Tuomala (1990) and implies that an average individual spends half of his/her time endowment working in the absence of taxation. (ii) The Gini-coefficients of first- and second-period consumption with binding credit constraints are around 0.32. OECD (2014b) shows that the OECD average of Gini-coefficients of income after taxes and transfers is about 0.32 in 2010. (iii) The share of credit-constrained individuals is approximately 20 % of the population, based on evidence provided in Hall and Mishkin (1982), Mariger (1986), and Jappelli (1990). The parameter values used for the benchmark simulation are summarized in Table 1.

We simulate the model for three important cases: (i) *Baseline*: this case assumes binding credit constraints and a redistributive government, (ii) *Perfect capital markets*: in this special case all credit constraints are slack, but the government is still redistributive, and (iii) *Pure efficiency*: this special case assumes binding credit constraints and a government that is non-redistributive, i.e., it maximizes the utility of a representative

[16] The OECD average of the marginal tax rates for a single individual earning 100 % of the average wage is 43.5 % in 2010 (OECD 2011). We compute effective marginal tax rates by correcting the marginal tax rate on earnings for indirect taxes. The OECD average of indirect tax rates is 10.4 % in 2010 (OECD 2011). We calculate the effective marginal tax rate as $\frac{t+\tau}{1+\tau} = 48.8\,\%$, where τ is the indirect tax rate.

[17] The calculation of the effective subsidy rate corrects for the different tax treatment of forgone earnings and direct costs of education. Forgone earnings are taxed at the income tax rate of 48.8 %. The share of direct costs in the total costs is around 0.25 (cf. Trostel 1993). The OECD average of subsidies on the direct costs of education is around 80 % (OECD 2014a). We thus calculate the effective subsidy rate s as $s = 1-(1-t)(1-\kappa)-(1-s)\kappa = 56.6\,\%$ where κ is the share of direct costs in total costs of education.

Table 1 Parameter values for benchmark simulation

$\varepsilon = 0.25$	Frisch elasticity of labor supply
$\beta = 0.5$	Elasticity of education investment
$\sigma = 0.67$	Intertemporal elasticity of substitution
$\zeta = 0$	Elasticity of inequality aversion
$\alpha = 0.42$	Share parameter of first-period consumption
$\nu = 0.25$	Income effect parameter
$r = 1.09$	Interest rate
$\gamma = 4.3$	Labor dis-utility parameter
$\mu_n = 3$	Mean of log-ability
$\sigma_n = 0.55$	Standard deviation of log-ability
$\mu_\omega = 3.5$	Mean of log-wealth
$\sigma_\omega = 0.7$	Standard deviation of log-wealth
$\varrho = 0.3$	Correlation coefficient

Table 2 Simulation results for the *Baseline*, *Perfect capital markets* and *Pure efficiency* cases

	Baseline	Perfect market	Pure efficiency
t (%)	56.6	36.2	43.6
s (%)	60.3	36.2	50.5
g	3.5	3.2	2.6
$\overline{\pi}$ (%)	8.4	0	12.3
$\overline{\Delta}$ (%)	−8.7	0	−12.3
Σ (%)	13.7	0	19.5
$\overline{e}$	9.84	13.1	10.5
$\overline{l}$	0.42	0.47	0.45
$\overline{a}$	15.01	8.34	12.8
$\overline{c^1}$	25.5	27.4	25.5
Gini (c^1)	0.31	0.31	0.32
$\overline{c^2}$	52.7	55.5	54.0
Gini (c^2)	0.31	0.31	0.32

All parameters of the model take the values provided in Table 1

individual.[18] These special cases allow us to fully trace down the main mechanisms of our model.

Table 2 provides the simulation results for optimal taxes t, subsidies s, transfers g, the income-weighted average of implicit taxes due to credit constraints $\overline{\pi}$, and the income-weighted average net tax on education $\overline{\Delta}$.[19] The shares of credit-constrained

[18] The objective of the government is to maximize $u\left(\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} c^1_{n\omega}\mathrm{d}F(n,\omega), \int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} c^2_{n\omega}\mathrm{d}F(n,\omega)\right) - v\left(\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty} l_{n\omega}\mathrm{d}F(n,\omega)\right)$.

[19] The programs are written in Mathematica and are available upon request.

individuals are denoted by Σ. The table provides the average levels of educational investment, labor supply, savings, and consumption levels in both periods. We also report the Gini-coefficients for consumption.

Our main theoretical result is that optimal taxes with credit constraints are higher than optimal taxes without credit constraints. Our simulations confirm this: The optimal tax rate of 56.6 % in the baseline with credit constraints is 55.2 % (20.4 %-points) higher than the optimal tax rate of 36.2 % under perfect capital markets. Moreover, our simulations confirm our theoretical result that, with binding credit constraints, education should be subsidized on a net basis: The average net tax on education $\bar{\Delta}$ is negative. Given that the implicit tax on human capital is non-negative ($\bar{\pi} \geq 0$), education subsidies s are thus always larger than income tax rates t. Moreover, we confirm Bovenberg and Jacobs (2005) that education is optimally subsidized at the same rate as the income tax (i.e., $s = t$) if there are no credit constraints.

When capital markets are perfect, 19.4 % of the population needs to borrow funds to finance education and first-period consumption. If borrowing is not possible at all, these people would be credit- constrained. Aggregate saving is therefore much lower when capital markets are perfect. Optimal tax and education policies reduce the share of credit-constrained individuals to 13.7 % of the population. Our Gini-measures for income inequality do not differ much between the cases with perfect capital markets and the baseline. On the one hand, credit constraints increase income inequality, due to the positive correlation between initial wealth and ability. On the other hand, the higher income tax rate when credit constraints are present corrects this adverse effect on inequality.

The pure efficiency case gives more evidence that optimal income taxes should be positive to alleviate credit constraints: Taxes are as high as 43.6 % even if the government is not interested in income redistribution at all. As in the baseline case, education is subsidized at a higher rate (i.e., $s > t$) and the net tax on education is negative ($\bar{\Delta} < 0$). Due to the absence of distributional concerns, the government sets income taxes and subsidies education at lower rates than in the baseline. Consequently, credit constraints are alleviated to a lesser degree. The number of credit-constrained individuals rises from 13.7 to 19.5 %. As a result, the implicit tax due to credit constraints $\bar{\pi}$ increases from 8 % in the baseline to 12.4 % in the pure efficiency case.

Our simulations show that the average net subsidy rate on education is almost equal to the average implicit tax on education, i.e., $\bar{\Delta} \approx -\bar{\pi}$, in both the baseline and the pure efficiency cases. This finding suggests that the second term in Eq. (38) is approximately zero. Hence, the welfare loss due to the regressive incidence of the net subsidy on education roughly cancels against the welfare gain of lower labor-supply distortions. Therefore, the optimal net tax on education is driven primarily by pure efficiency reasons: to alleviate explicit and implicit distortions on human capital formation and to alleviate credit constraints.

We carried out various robustness checks by varying the main elasticities of our model: σ, ε and β. Table 3 provides the optimal tax rates where we compare the optimal policies with credit constraints to those obtained under perfect capital markets and in the pure efficiency case. We also report the income-weighted average implicit tax on human capital $\overline{\pi}$ and the income-weighted average net tax on education $\overline{\Delta}$.

Table 3 Simulation results optimal tax policy for different values for the elasticity of substitution, the Frisch elasticity of labor supply, and the elasticity of human capital

	Baseline					Perfect market			Pure efficiency				
	$t\,(\%)$	$s\,(\%)$	g	$\overline{\pi}\,(\%)$	$\overline{\Delta}\,(\%)$	$t\,(\%)$	$s\,(\%)$	g	$t\,(\%)$	$s\,(\%)$	g	$\overline{\pi}\,(\%)$	$\overline{\Delta}\,(\%)$
$\sigma = 0.25$	59.3	62.5	3.6	7.9	−7.7	36.3	36.3	3.3	47.3	53.6	2.8	12.0	−12.0
$\sigma = 1.25$	54.7	58.6	3.5	8.7	−8.4	36.2	36.2	3.2	40.3	48.4	2.4	13.5	−13.5
$\sigma = 2.0$	52.7	56.9	3.4	8.8	−9.0	36.2	36.2	3.2	35.0	44.4	2.0	14.6	−14.6
$\varepsilon = 0.1$	74.7	76.3	10.0	6.3	−6.5	54.0	54.0	8.6	68.2	71.3	9.1	9.7	−9.7
$\varepsilon = 0.15$	68.4	70.7	7.0	7.5	−7.5	46.4	46.4	6.1	58.8	63.5	5.9	11.5	−11.5
$\varepsilon = 0.3$	51.6	55.6	2.5	8.3	−8.2	32.4	32.4	2.4	35.4	43.6	1.6	12.7	−12.7
$\beta = 0.3$	27.6	28.5	0.8	1.3	−1.3	24.0	24.0	0.7	1.3	4.5	0.0	3.3	−3.3
$\beta = 0.4$	38.4	40.6	1.5	3.6	−3.6	28.5	28.5	1.3	16.8	22.7	0.6	7.1	−7.1
$\beta = 0.6$	79.3	82.6	11.2	16.0	−15.8	47.2	47.2	15.7	75.6	80.2	10.5	19.1	−19.0

All parameters of the model take the values provided in Table 1, except where indicated otherwise

Table 3 shows that the optimal tax rates with credit constraints are always higher than those with perfect capital markets and the pure efficiency case. Moreover, the optimal subsidy rates are always slightly higher than the optimal tax rates and education is (on average) always subsidized on a net basis when credit constraints are present, i.e., $\bar{\Delta} < 0$. Note that in the case of perfect capital markets, the implicit tax on human capital is zero $\bar{\pi} = 0$, and education is not taxed on a net basis, i.e., $\bar{\Delta} = 0$ and $t = s$ (cf. Bovenberg and Jacobs 2005). Positive optimal tax rates are found in the pure efficiency case as well.

In the baseline and in the pure efficiency case, the optimal income tax rate decreases when the intertemporal elasticity of substitution in consumption is increased, i.e., σ is larger. Intuitively, credit constraints become less severe if individuals are more willing to substitute first-period consumption for second-period consumption. Under perfect capital markets, the optimal tax rates remain the same irrespective of the value for σ. Human capital investments are independent from consumption choices when capital markets are perfect.

Optimal tax rates decrease with a higher elasticity of labor supply ε. Taxes are more distortionary when labor supply becomes more elastic. When the elasticity of the human capital production β increases, the baseline optimal taxes significantly increase. There are three counter-acting forces at play here. First, the optimal tax rate decreases because taxes are more distortionary—for a given desire to redistribute income and to alleviate credit constraints. Second, the optimal tax rate increases since there is more income inequality when β increases—for given distortions. Third, returns to education rise if β increases. Hence, individuals want to invest more in human capital and credit constraints become tighter, as reflected by a larger average implicit tax rate $\bar{\pi}$. Consequently, optimal tax rates increase. Clearly, the second and/or third effects always dominate the first effect. In particular, in the pure efficiency case the second effect is absent, since the government does not want to levy income taxes for redistributional concerns. In the case with perfect capital markets the third effect is absent, since the government does not need to raise taxes to alleviate credit constraints. Indeed, if $\beta = 0.3$, credit constraints are the least severe (i.e., the average implicit tax on human capital $\bar{\pi}$ is lowest) and the optimal income tax rate in the pure efficiency case is close to zero. In this case, the baseline optimal tax rate of 27.6 % is driven mainly by redistributional concerns.

6 Conclusion

This paper has formulated a two-period life cycle model of saving, labor supply, and human capital investment when individuals differ in their ability and their initial wealth to analyze optimal income taxes and education subsidies. Binding borrowing constraints cause sub-optimal smoothing of consumption over the life cycle and sub-optimal investment in human capital. We have demonstrated that the optimal linear income tax is always positive—even in the absence of redistributional concerns. A distortionary income tax is optimal because it relaxes borrowing constraints by redistributing resources from the unconstrained to the borrowing constrained stages of the life cycle. Hence, redistribution allows for better consumption smoothing and larger investments in human capital. The equity-efficiency trade-off is therefore less severe when progressive income taxes mitigate capital-market imperfections, as the progressive income tax helps to correct a non-tax distortion in the capital market.

We allow for subsidies on human capital investment to study how the optimal mix of income taxes and education subsidies is affected by binding credit constraints. In the absence of redistributional concerns, human capital should be subsidized on a net basis to relax credit constraints, to alleviate distortions on human capital from income taxes, and to reduce distortions on human capital from implicit taxes that originate from capital-market imperfections. When redistributional concerns are allowed for, human capital might be subsidized even more if this helps to reduce tax distortions in labor supply. However, the optimal education subsidy could also be lower, since education subsidies are regressive. Whether education is taxed or subsidized on a net basis therefore becomes theoretically ambiguous.

We simulated a realistically calibrated version of our model to quantitatively explore the relevance of capital-market imperfections for the setting of optimal income taxes and education subsidies. Simulated optimal income taxes are around 57 % when capital markets fail, which is about 55 % (20 %-points) higher than in the case where capital markets are perfect. Optimal income taxes are even as high as 44 % for pure efficiency reasons, i.e., in the absence of any distributional concerns. Moreover, education is subsidized on a net basis. If capital-market imperfections are more important, education subsidies are optimally higher. Education subsidies are shown to primarily serve an efficiency role as the benefits of education subsidies to lower labor-supply distortions largely cancel against the cost of their regressive incidence.

Acknowledgments The authors thank the editor, Andreas Haufler, two anonymous referees, Dan Anderberg, Dirk Schindler, Bernd Genser, Mathias Kifmann and seminar participants at the University of Konstanz, the International Public Finance Conference in Cape Town, the Association for Public Economic Theory Conference in Galway, and the European Economic Association Conference in Barcelona for useful comments and suggestions. All remaining errors are our own. Bas Jacobs gratefully acknowledges financial support from the Dutch Organisation for Sciences (NWO Vidi Grant No. 452-07-013, "Skill Formation in Distorted Labor Markets") and Hongyan Yang acknowledges financial support from the German Research Council (CRC 597: transformations of the State).

Appendix A: Optimal taxation

We derive the optimal tax formulae for the case with heterogeneous agents. The representative-agent case is a special case where the distributional characteristics are set to zero and there is no income-weighting of elasticities and social costs of credit constraints. The Lagrangian for maximizing social welfare is given by:

$$\mathcal{L} \equiv \int_{\underline{n}}^{\infty} \int_{\underline{\omega}}^{\infty} [\Psi(V_{n\omega}) + \eta[tnl_{n\omega}\phi_{n\omega} - (2+r)g - s(1+r)e_{n\omega}] \, \mathrm{d}F(n, \omega). \quad (42)$$

where η is the Lagrange multiplier on the government budget constraint. First-order conditions for the lump-sum transfers, income taxes, and education subsidies are, respectively:

$$\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[\Psi'(V_{n\omega})(u_{1,n\omega}+u_{2,n\omega})-(2+r)\right]\mathrm{d}F(n,\omega)$$
$$+\eta\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[tn\phi_{n\omega}\frac{\partial l_{n\omega}}{\partial g}+\left(tnl_{n\omega}\phi'_{n\omega}-s(1+r)\right)\frac{\partial e_{n\omega}}{\partial g}\right]\mathrm{d}F(n,\omega)=0, \tag{43}$$

$$\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[-\Psi'(V_{n\omega})u_{2,n\omega}nl_{n\omega}\phi_{n\omega}+\eta nl_{n\omega}\phi_{n\omega}\right]\mathrm{d}F(n,\omega)$$
$$+\eta\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[tn\phi_{n\omega}\frac{\partial l_{n\omega}}{\partial t}+\left(tnl_{n\omega}\phi'_{n\omega}-s(1+r)\right)\frac{\partial e_{n\omega}}{\partial t}\right]\mathrm{d}F(n,\omega)=0, \tag{44}$$

$$\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[\Psi'(V_{n\omega})u_{1,n\omega}e_{n\omega}-\eta(1+r)e_{n\omega}\right]\mathrm{d}F(n,\omega)$$
$$+\eta\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[tn\phi_{n\omega}\frac{\partial l_{n\omega}}{\partial s}+\left(tnl_{n\omega}\phi'_{n\omega}-s(1+r)\right)\frac{\partial e_{n\omega}}{\partial s}\right]\mathrm{d}F(n,\omega)=0, \tag{45}$$

where we used Roy's identity in each expression: $\frac{\partial V_{n\omega}}{\partial g}=u_{1,n\omega}+u_{2,n\omega}$, $\frac{\partial V_{n\omega}}{\partial t}=-u_{2,n\omega}nl_{n\omega}\phi_{n\omega}$, and $\frac{\partial V_{n\omega}}{\partial s}=u_{1,n\omega}e_{n\omega}$.

A.1 Optimal lump-sum transfer

Define the social marginal value of income of an individual with ability n and initial wealth ω as

$$b_{n\omega}\equiv\frac{\Psi'(V_{n\omega})\left(u_{1,n\omega}+u_{2,n\omega}\right)}{\eta}+tn\phi_{n\omega}\frac{\partial l_{n\omega}}{\partial g}+\left(tnl_{n\omega}\phi'_{n\omega}-s(1+r)\right)\frac{\partial e_{n\omega}}{\partial g}. \tag{46}$$

Use (46) in Eq. (43) to find

$$\bar{b}=2+r. \tag{47}$$

A.2 Optimal labor tax

Substitute the Slutsky' equations (17) and (18) in (44), use the definitions (46) and $z_{n\omega}\equiv nl_{n\omega}\phi_{n\omega}$, and rearrange to derive:

$$\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[z_{n\omega}-b_{n\omega}\frac{u_{2,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}z_{n\omega}+tn\phi_{n\omega}\frac{\partial l^c_{n\omega}}{\partial t}\right.$$
$$\left.+\left(tnl_{n\omega}\phi'_{n\omega}-s(1+r)\right)\frac{\partial e^c_{n\omega}}{\partial t}\right]\mathrm{d}F(n,\omega)=0. \tag{48}$$

Next, define the compensated tax elasticities as follows:

$$\varepsilon_{lt,n\omega}\equiv-\frac{\partial l^c_{n\omega}}{\partial t}\frac{(1-t)}{l_{n\omega}},\quad \varepsilon_{et,n\omega}\equiv-\frac{\partial c^c_{n\omega}}{\partial t}\frac{(1-t)}{l_{n\omega}}. \tag{49}$$

Substitute (49) in (48) to get:

$$\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[z_{n\omega}-\frac{u_{2,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}b_{n\omega}z_{n\omega}\right]\mathrm{d}F(n,\omega)$$
$$-\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[\frac{t}{1-t}\varepsilon_{lt,n\omega}z_{n\omega}+\frac{\left(tnl_{n\omega}\phi'_{n\omega}-s(1+r)\right)}{1-t}\varepsilon_{et,n\omega}e_{n\omega}\right]\mathrm{d}F(n,\omega)=0. \tag{50}$$

Use the first-order condition for learning (14) twice to derive:

$$\frac{\left(tnl_{n\omega}\phi'_{n\omega}-s(1+r)\right)}{1-t}e_{n\omega}=\frac{t}{1-t}\beta z_{n\omega}-\frac{s}{1-s}(1-\pi_{n\omega})\beta z_{n\omega}. \tag{51}$$

Substituting (51) in (50), dividing by $\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)$, and rearranging gives:

$$\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[z_{n\omega}-\frac{u_{2,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}b_{n\omega}z_{n\omega}\right]\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}=\frac{t}{1-t}\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\varepsilon_{lt,n\omega}z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}$$
$$+\frac{t}{1-t}\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\beta\varepsilon_{et,n\omega}z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}-\frac{s}{1-s}\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}(1-\pi_{n\omega})\beta\varepsilon_{et,n\omega}z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}. \tag{52}$$

Substitute the definition for $\pi_{n\omega}$ to find $\frac{u_{2,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}=\frac{1-\pi_{n\omega}}{2+r-\pi_{n\omega}}$, use the definition for $\chi_{n\omega}\equiv 1-\frac{1+r}{2+r-\pi_{n\omega}}$ and $\bar{b}\equiv\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}b_{n\omega}\mathrm{d}F(n,\omega)=2+r$, and substitute for (33) to rewrite the left-hand side of (52) as:

$$\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[z_{n\omega}-\frac{u_{2,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}b_{n\omega}z_{n\omega}\right]\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}$$
$$=\xi_z+\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\pi_{n\omega}(1-\chi_{n\omega})\frac{b_{n\omega}}{2+r}z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)\bar{b}}. \tag{53}$$

Note that we added and subtracted $\frac{b_{n\omega}}{2+r}z_{n\omega}$ in the numerator to derive the equation. Consequently, by substituting (53) in (52), the optimal income tax can be written as:

$$\frac{t}{1-t}(\overline{\varepsilon_{lt}}+\beta\overline{\varepsilon_{et}})-\frac{s}{1-s}\beta\overline{(1-\pi)\varepsilon_{et}}=\xi_z+\frac{\overline{(1-\chi)\pi b}}{\bar{b}}. \tag{54}$$

where all bars indicate income-weighted variables, except for the social value of income $\bar{b}$. The representative-agent case follows by setting $\xi_z=0$ and dropping all the bars over income-weighted variables:

$$\frac{t}{1-t}(\varepsilon_{lt}+\beta\varepsilon_{et})-\frac{s}{1-s}(1-\pi)\beta\varepsilon_{et}=(1-\chi)\pi. \tag{55}$$

A.3 Optimal education subsidy

Substitute the Slutsky equations (19) and (20) in (45), use the definition (46), and rearrange to derive:

$$\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[-(1+r)e_{n\omega}+b_{n\omega}\frac{u_{1,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}e_{n\omega}\right]\mathrm{d}F(n,\omega)$$
$$+\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[tn\phi_{n\omega}\frac{\partial l^c_{n\omega}}{\partial s}+\left(tnl_{n\omega}\phi'_{n\omega}-s(1+r)\right)\frac{\partial e^c_{n\omega}}{\partial s}\right]\mathrm{d}F(n,\omega)=0. \tag{56}$$

Next, define the compensated subsidy elasticities as follows:

$$\varepsilon_{ls,n\omega}\equiv\frac{\partial l^c_{n\omega}}{\partial s}\frac{(1-s)}{l_{n\omega}},\quad \varepsilon_{es,n\omega}\equiv\frac{\partial e^c_{n\omega}}{\partial s}\frac{(1-s)}{e_{n\omega}}. \tag{57}$$

Substitute (51) and (57) in (56) and divide by $\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)$ to find:

$$-\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[-(1+r)e_{n\omega}+b_{n\omega}\frac{u_{1,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}e_{n\omega}\right]\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}$$
$$=\frac{t}{1-s}\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\varepsilon_{ls,n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}$$
$$+\frac{t}{1-s}\beta\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\varepsilon_{es,n\omega}z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}$$
$$-\frac{s}{1-s}\frac{(1-t)}{(1-s)}\beta\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}(1-\pi_{n\omega})\varepsilon_{es,n\omega}z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}. \tag{58}$$

Use the definition for $\pi_{n\omega}$ to find $\frac{u_{2,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}=\frac{1-\pi_{n\omega}}{2+r-\pi_{n\omega}}$, use the definition for $\chi_{n\omega}\equiv 1-\frac{1+r}{2+r-\pi_{n\omega}}$ and substitute (34), (51) to rewrite the left-hand side of (58) as:

$$\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left[-(1+r)e_{n\omega}+b_{n\omega}\frac{u_{1,n\omega}}{u_{1,n\omega}+u_{2,n\omega}}e_{n\omega}\right]\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}\frac{(1-s)}{(1-t)}$$
$$=-\xi_e\beta\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}(1-\pi_{n\omega})z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)}$$
$$+\beta\frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\chi_{n\omega}\pi_{n\omega}b_{n\omega}z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\mathrm{d}F(n,\omega)\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}b_{n\omega}\mathrm{d}F(n,\omega)}. \tag{59}$$

Note that we added and subtracted $\left(\frac{1+r}{2+r}\right) b_{n\omega}e_{n\omega}$ in the numerator to derive the equation. By substituting (59) in (58), we find the optimal education subsidy:

$$\frac{t}{1-t}\frac{\overline{\varepsilon_{ls}}}{\beta}+\frac{t}{1-t}\overline{\varepsilon_{es}}-\frac{s}{1-s}\overline{(1-\pi)\varepsilon_{es}}=\xi_e\overline{(1-\pi)}-\frac{\overline{\chi\pi b}}{\bar{b}}. \tag{60}$$

The representative-agent case follows by setting $\xi_e=0$ and dropping all bars:

$$\frac{t}{1-t}\frac{\varepsilon_{ls}}{\beta}+\frac{t}{1-t}\varepsilon_{es}-\frac{s}{1-s}(1-\pi)\varepsilon_{es}=-\chi\pi. \tag{61}$$

Appendix B: Optimal tax structure

The optimal income tax and education subsidy are given by (54) and (60). This is a system of two equations in the optimal tax rate t and optimal subsidy rate s. Solving for the optimal tax system gives:

$$\frac{t}{1-t}=\frac{\left(\xi_z+\frac{\overline{(1-\chi)\pi b}}{\bar{b}}\right)\overline{(1-\pi)\varepsilon_{es}}-\left(\xi_e\overline{(1-\pi)}-\frac{\overline{\chi\pi b}}{\bar{b}}\right)\beta\overline{(1-\pi)\varepsilon_{et}}}{(\overline{\varepsilon_{lt}}+\beta\overline{\varepsilon_{et}})\overline{(1-\pi)\varepsilon_{es}}-\left(\frac{\overline{\varepsilon_{ls}}}{\beta}+\overline{\varepsilon_{es}}\right)\beta\overline{(1-\pi)\varepsilon_{et}}}, \tag{62}$$

$$\frac{s}{1-s}=\frac{\left(\xi_z+\frac{\overline{(1-\chi)\pi b}}{\bar{b}}\right)\left(\frac{\overline{\varepsilon_{ls}}}{\beta}+\overline{\varepsilon_{es}}\right)-\left(\xi_e\overline{(1-\pi)}-\frac{\overline{\chi\pi b}}{\bar{b}}\right)(\overline{\varepsilon_{lt}}+\beta\overline{\varepsilon_{et}})}{(\overline{\varepsilon_{lt}}+\beta\overline{\varepsilon_{et}})\overline{(1-\pi)\varepsilon_{es}}-\left(\frac{\overline{\varepsilon_{ls}}}{\beta}+\overline{\varepsilon_{es}}\right)\beta\overline{(1-\pi)\varepsilon_{et}}}. \tag{63}$$

The representative-agent case follows upon setting the distributional characteristics to zero ($\xi_z=\xi_e=0$) and dropping all the bars:

$$\frac{t}{1-t}=\left(\frac{(1-\chi)\varepsilon_{es}+\chi\beta\varepsilon_{et}}{\varepsilon_{lt}\varepsilon_{es}-\varepsilon_{ls}\varepsilon_{et}}\right)\pi, \tag{64}$$

$$\frac{(1-\pi)s}{1-s}=\left(\frac{(1-\chi)\left(\frac{\varepsilon_{ls}}{\beta}+\varepsilon_{es}\right)+\chi(\varepsilon_{lt}+\beta\varepsilon_{et})}{\varepsilon_{lt}\varepsilon_{es}-\varepsilon_{ls}\varepsilon_{et}}\right)\pi. \tag{65}$$

The first-order conditions for t and s also yield measures for the average net tax on education $\bar{\Delta}$. In order to derive an analytically tractable solution, we assume that the correlation between credit constraints and behavioral elasticities are zero so as to obtain $\frac{\overline{(1-\pi)\varepsilon_{et}}}{\overline{(1-\pi)}\overline{\varepsilon_{et}}}=1$, $\frac{\overline{(1-\pi)\varepsilon_{es}}}{\overline{(1-\pi)}\overline{\varepsilon_{es}}}=1$ and $\frac{\overline{(1-\pi)\varepsilon_{et}}}{\overline{\varepsilon_{et}}}=\overline{(1-\pi)}$. Consequently, we have

$$\bar{\Delta}=\frac{t}{1-t}-\frac{\overline{(1-\pi)}s}{1-s}=-\left(\frac{\frac{\overline{\chi\pi b}}{\bar{b}}\overline{\varepsilon_{lt}}+\frac{\overline{(1-\chi)\pi b}}{\bar{b}}\frac{\overline{\varepsilon_{ls}}}{\beta}+\xi_z\frac{\overline{\varepsilon_{ls}}}{\beta}-\xi_e\overline{(1-\pi)}\overline{\varepsilon_{lt}}}{\overline{\varepsilon_{lt}}\;\overline{\varepsilon_{es}}-\overline{\varepsilon_{ls}}\;\overline{\varepsilon_{et}}}\right). \tag{66}$$

Rewriting yields the expression in the main text. In the case of a representative agent, we set the distributional characteristics to zero ($\xi_z = \xi_e = 0$) and drop the bars to obtain

$$\Delta = -\left(\frac{\varepsilon_{lt}\chi + \frac{\varepsilon_{ls}}{\beta}(1-\chi)}{\varepsilon_{lt}\varepsilon_{es} - \varepsilon_{ls}\varepsilon_{et}}\right)\pi. \tag{67}$$

Rewriting yields the expression in the main text.

Appendix C: Optimal tax and education policy with perfect capital markets

As a special case, our model harbors the model of Bovenberg and Jacobs (2005). In particular, with perfect capital markets we have $\pi_{n\omega} = 0$, $\frac{u_{1,n\omega}}{u_{2,n\omega}} = 1 + r$, and $1 - \chi_{n\omega} = \frac{1+r}{2+r}$. Note furthermore from (14) that there is a linear relationship between $z_{n\omega}$ and $e_{n\omega}$: $e_{n\omega} = \frac{\beta(1-t)}{(1+r)(1-s)} z_{n\omega}$. Substitution of this relationship in the distributional characteristic for education (34) gives:

$$\begin{aligned}
\xi_e &\equiv \frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left(1-\frac{b_{n\omega}}{2+r}\right)e_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}e_{n\omega}\mathrm{d}F(n,\omega)\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{b_{n\omega}}{2+r}\mathrm{d}F(n,\omega)} \\
&= \frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left(1-\frac{b_{n\omega}}{2+r}\right)\frac{\beta(1-t)}{(1+r)(1-s)}z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\beta(1-t)}{(1+r)(1-s)}z_{n\omega}\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{b_{n\omega}}{2+r}\mathrm{d}F(n,\omega)} \\
&= \frac{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\left(1-\frac{b_{n\omega}}{2+r}\right)z_{n\omega}\mathrm{d}F(n,\omega)}{\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}z_{n\omega}\int_{\underline{n}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{b_{n\omega}}{2+r}\mathrm{d}F(n,\omega)} \equiv \xi_z.
\end{aligned} \tag{68}$$

Substituting all these relationships in first-order conditions for optimal taxes (54) and education subsidies (60) then gives (using $\xi \equiv \xi_z = \xi_e$):

$$\xi = \frac{t}{1-t}\overline{\varepsilon_{lt}} + \left(\frac{t}{1-t} - \frac{s}{1-s}\right)\beta\overline{\varepsilon_{et}}, \tag{69}$$

$$\xi = \frac{t}{1-t}\frac{\overline{\varepsilon_{ls}}}{\beta} + \left(\frac{t}{1-t} - \frac{s}{1-s}\right)\overline{\varepsilon_{es}}. \tag{70}$$

Subtracting both equations yields

$$\begin{aligned}
0 &= \frac{t}{1-t}\left(\overline{\varepsilon_{lt}} - \frac{\overline{\varepsilon_{ls}}}{\beta}\right) + \left(\frac{t}{1-t} - \frac{s}{1-s}\right)(\beta\overline{\varepsilon_{et}} - \overline{\varepsilon_{es}}) \\
&= \left(\frac{t}{1-t} - \frac{s}{1-s}\right)(\beta\overline{\varepsilon_{et}} - \overline{\varepsilon_{es}}),
\end{aligned} \tag{71}$$

where we used $\overline{\varepsilon_{lt}} = \overline{\varepsilon_{ls}}/\beta$ (see online Appendix C). Consequently, since $\beta\overline{\varepsilon_{et}} \neq \overline{\varepsilon_{es}}$, see online Appendix C, the optimal policy is given by $s = t$ and $\frac{t}{1-t} = \xi/\overline{\varepsilon_{lt}}$.

References

Acemoglu, D., & Pischke, J. S. (2001). Changes in the wage structure, family income, and children's education. *European Economic Review*, *45*(4–6), 890–904.

Aiyagari, S. R. (1995). Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy*, *103*(6), 1158–1175.

Akerlof, G. (1976). The economics of caste and of the rat race and other woeful tales. *Quarterly Journal of Economics*, *90*(4), 599–617.

Anderberg, D. (2009). Optimal policy and the risk-properties of human capital reconsidered. *Journal of Public Economics*, *93*(9–10), 1017–1026.

Anderberg, D., & Andersson, F. (2003). Investments in human capital, wage uncertainty, and public policy. *Journal of Public Economics*, *87*(7–8), 1521–1537.

Andolfatto, D., & Gervais, M. (2006). Human capital investment and debt constraints. *Review of Economic Dynamics*, *9*(1), 52–67.

Atkinson, A. B., & Stiglitz, J. E. (1980). *Lectures on public economics*. New York: McGraw-Hill.

Attanasio, O. P., & Weber, G. (2010). Consumption and saving: Models of intertemporal allocation and their implications for public policy. *Journal of Economic Literature*, *48*(3), 693–751.

Belley, P., & Lochner, L. (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human Capital*, *1*(1), 37–89.

Benabou, R. (1996a). Heterogeneity, stratification, and growth: Macroeconomic implications of community structure and school finance. *The American Economic Review*, *86*(3), 584–609.

Benabou, R. (1996b). Equity and efficiency in human capital investment: The local connection. *Review of Economic Studies*, *63*(2), 237–264.

Blundell, R., & MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics, Chap 27* (Vol. 3A, pp. 1559–1695). North Holland: Amsterdam.

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

Boone, J., & Bovenberg, A. L. (2002). Optimal labour taxation and search. *Journal of Public Economics*, *85*(1), 53–97.

Bovenberg, A. L. (2006). Tax policy and labor market performance. In J. Agell & P. Sørensen (Eds.), *Tax policy and labour market performance* (pp. 3–74). Cambridge, MA: CESifo and MIT Press.

Bovenberg, A. L., & Jacobs, B. (2005). Redistribution and education subsidies are Siamese twins. *Journal of Public Economics*, *89*(11–12), 2005–2035.

Chamley, C. (2001). Capital income taxation, wealth distribution and borrowing constraints. *Journal of Public Economics*, *79*(1), 55–69.

Caneiro, P., & Heckman, J. J. (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, *112*(482), 705–734.

De Gregorio, J. (1996). Borrowing constraints, human capital accumulation, and growth. *Journal of Monetary Economics*, *37*(1), 49–71.

Diamond, P. A. (1975). A many-person Ramsey tax rule. *Journal of Public Economics*, *4*(4), 335–342.

Durlauf, S. N. (1996). A theory of persistent income inequality. *Journal of Economic Growth*, *1*(1), 75–93.

Eaton, J., & Rosen, H. S. (1980). Taxation, human capital, and uncertainty. *American Economic Review*, *70*(4), 705–715.

Eurostat, (2014). *Taxation trends in the European Union*. Luxembourg: European Union.

Evers, M., de Mooij, R. A., & van Ruuren, D. (2008). The wage elasticity of labour supply: A synthesis of empirical estimates. *De Economist*, *156*(1), 26–43.

Fernandez, R., & Rogerson, R. (1996). Income distribution, communities, and the quality of public education. *The Quarterly Journal of Economics*, *111*(1), 135–164.

Fernandez, R., & Rogerson, R. (1998). Public education and income distribution: A dynamic quantitative evaluation of education-finance reform. *The American Economic Review*, *88*(4), 813–833.

Findeisen, S., & Sachs, D. (2013). Education and optimal dynamic taxation: The role of income-contingent student loans, Working Paper, University of Mannheim/Cologne.

Galor, O., & Moav, O. (2004). From physical to human capital accumulation: Inequality an the process of development. *Review of Economic Studies*, *71*(4), 1001–1026.

Galor, O., & Zeira, J. (1993). Income distribution and macroeconomics. *Review of Economic Studies*, *60*(1), 35–52.

Glomm, G., & Ravikumar, B. (1992). Public versus private investment in human capital: Endogenous growth and income inequality. *Journal of Political Economy*, *100*(4), 818–834.

Guvenen, F. (2006). Reconciling conflicting evidence on the elasticity of intertemporal substitution: A macroeconomic perspective. *Journal of Monetary Economics*, *53*(7), 1451–1472.

Hall, R. E., & Mishkin, F. S. (1982). The sensitivity of consumption to transitory income: Estimates from panel data on households. *Econometrica*, *50*(2), 461–481.

Hamilton, J. (1987). Optimal wage and income taxation with wage uncertainty. *International Economic Review*, *28*(2), 373–388.

Hare, P. G., & Ulph, D. T. (1979). On education and distribution. *Journal of Political Economy*, *87*(5) pt 2, 193–212.

Hoff, K., & Lyon, A. B. (1995). Non-leaky buckets: Optimal redistributive taxation and agency costs. *Journal of Public Economics*, *58*(3), 365–390.

Hosios, A. J. (1990). On the efficiency of matching and related models of search and unemployment. *Review of Economic Studies*, *57*(2), 279–298.

Hubbard, R. G., & Judd, K. L. (1986). Liquidity constraints, fiscal policy, and consumption. *Brookings Papers on Economic Activity*, *1*(1), 1–51.

İmrohoroğlu, S. (1998). A quantitative analysis of capital income taxation. *International Economic Review*, *39*(2), 307–328.

Jacobs, B. (2005). Optimal income taxation with endogenous human capital. *Journal of Public Economic Theory*, *7*(2), 295–315.

Jacobs, B. (2012). Optimal redistributive tax and education policies in general equilibrium. *International Tax and Public Finance*, *20*(2), 312–337.

Jacobs, B., & Bovenberg, A. L. (2008). Optimal taxation of human capital and the earnings function, CESifo working paper no. 2250, Munich: CESifo.

Jacobs, B., & Bovenberg, A. L. (2010). Human capital and optimal positive taxation of capital income. *International Tax and Public Finance*, *17*(5), 451–478.

Jacobs, B., & Bovenberg, A. L. (2011). Optimal taxation of human capital and the earnings function. *Journal of Public Economic Theory*, *13*(6), 957–971.

Jacobs, B., Schindler, D., & Yang, H. (2012). Optimal taxation of risky human capital. *Scandinavian Journal of Economics*, *114*(3), 908–931.

Jacobs, B., & van Wijnbergen, S. J. G. (2007). Capital-market failure, adverse selection, and equity financing of higher education. *FinanzArchiv*, *63*(1), 1–32.

Jacobs, B., & Yang, H. (2013). Second-best income taxation with endogenous human capital and borrowing constraints, CESifo working paper no. 4155, Munich: CESifo.

Jacquet, L., Lehmann, E., & Van der Linden, B. (2013). The optimal marginal tax rates with both intensive and extensive responses. *Journal of Economic Theory*, *148*(5), 1770–1805.

Jappelli, T. (1990). Who is credit constrained in the U.S. economy? *The Quarterly Journal of Economics*, *105*(1), 219–234.

Jones, L. E., Manuelli, R. E., & Rossi, P. E. (1993). Optimal taxation in models of endogenous growth. *Journal of Political Economy*, *101*(3), 485–517.

Judd, K. L. (1999). Optimal taxation and spending in general competitive growth models. *Journal of Public Economics*, *71*(1), 1–26.

Kane, T. J. (1995). *Rising public college tuition and college entry: How well do public subsidies promote access to college*. NBER working paper no. 5164. Cambridge, MA: NBER.

Kane, T. J. (1996). College cost, borrowing constraints and the timing of college entry. *Eastern Economic Journal*, *22*(2), 181–194.

Keane, M. P., & Wolpin, K. I. (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review*, *42*(4), 1051–1103.

Kleven, H. J., Kreiner, C. T., & Saez, E. (2009). The optimal income taxation of couples. *Econometrica*, *77*(2), 537–560.

Koskela, E., & Vilmunen, J. (1996). Tax progression is good for employment in popular models of trade union behaviour. *Labour Economics*, *3*(1), 65–80.

Krueger, D., & Ludwig, A. (2013). Optimal progressive labor income taxation and education subsidies when education decisions and intergenerational transfers are endogenous. *American Economic Review*, *103*(3), 496–501.

Layard, R. (1980). Human satisfactions and public policy. *The Economic Journal*, *90*(360), 737–750.

Layard, R. (2005). *Happiness: Lessons from a new science*. London: Penguin.

Lochner, L., & Monge-Naranjo, A. (2011). The nature of credit constraints and human capital. *American Economic Review*, *101*(6), 2487–2529.
Loury, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica*, *49*(4), 843–867.
Maldonado, D. (2008). Education policies and optimal taxation. *International Tax and Public Finance*, *15*(2), 131–143.
Mariger, R. P. (1986). *Consumption behavior and the effects of fiscal policies*. Cambridge-MA: Harvard University Press.
Mirrlees, J. A. (1971). An exploration in the theory of optimal income taxation. *Review of Economic Studies*, *38*(2), 175–208.
Mookherjee, D., & Ray, D. (2003). Persistent inequality. *Review of Economic Studies*, *70*(2), 369–393.
Nielsen, S. B., & Sørensen, P. B. (1997). On the optimality of the Nordic system of dual income taxation. *Journal of Public Economics*, *63*(3), 311–329.
OECD. (2011). *Taxation and employment, tax policy studies*, No. 21, Paris: OECD.
OECD. (2014a). *Education at a glance*. Paris: OECD.
OECD. (2014b). *Income distribution database*. Paris: OECD.
Pereira, Joana. (2009). *Essays on time-consistent fiscal policy*. Florence: European University Institute.
Piketty, T., & Saez, E. (2013). A theory of optimal inheritance taxation. *Econometrica*, *81*(5), 1851–1886.
Pissarides, C. A. (1998). The impact of employment tax cuts on unemployment and wages; the role of unemployment benefits and tax structure. *European Economic Review*, *42*(1), 155–183.
Plug, E. J. S., & Vijverberg, W. (2005). Does family income matter for schooling outcomes? Using adoptees as a natural experiment. *The Economic Journal*, *115*(506), 879–906.
Schindler, D. (2011). Tuition fees and the dual income tax—The optimality of the Nordic income tax system reconsidered. *German Economic Review*, *12*(1), 59–84.
Schindler, D., & Yang, H. (2015). Catalysts for social insurance: Education subsidies versus physical capital taxation. *International Tax and Public Finance*, *22*(2), 274–310.
Sørensen, P. B. (1999). Optimal tax progressivity in imperfect labor markets. *Labour Economics*, *6*(3), 435–452.
Stantcheva, S. (2014). Optimal taxation and human capital policies over the life cycle, working paper, Harvard University.
Stern, N. H. (1976). On the specifications of models of optimum income taxation. *Journal of Public Economics*, *6*(1–2), 123–162.
Stiglitz, J. E. (1994). *Whither Socialism?*. Cambridge, MA: MIT-Press.
Stinebrickner, R., & Stinebrickner, T. (2008). The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *American Economic Review*, *98*(5), 2163–2184.
Tobin, J. (1982). *Asset accumulation and economic activity: Reflections on contemporary macroeconomic theory*. University of Chicago Press.
Trostel, P. A. (1993). The effect of taxation on human capital. *Journal of Political Economy*, *101*(2), 327–350.
Tuomala, M. (1990). *On the optimal income tax and redistribution*. Oxford: Clarendon Press.
Ulph, D. (1977). On the optimal distribution of income and educational expenditure. *Journal of Public Economics*, *8*(3), 341–356.
Varian, H. R. (1980). Redistributive taxation as social insurance. *Journal of Public Economics*, *14*(1), 49–68.
Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review*, *43*(4), 1249–1287.
Van der Ploeg, F. (2006a). Are the welfare state and redistribution really that bad for the economy? Effects of reciprocal altruism, consumer rivalry and second-best. In F. Farina & E. Savaglio (Eds.), *Inequality and Economic Integration* (pp. 191–219). London: Routledge.
Van der Ploeg, F. (2006b). Do social policies harm employment? In J. Agell & P. Sørensen (Eds.), *Tax policy and labour market performance* (pp. 97–144). Cambridge, MA: CESifo and MIT Press.
Van Ewijk, C., & Tang, P. J. G. (2007). Unions, progressive taxes, and education subsidies. *European Journal of Political Economy*, *23*(4), 1119–1139.