Downloaded from UvA-DARE, the institutional repository of the University of Amsterdam (UvA)
http://hdl.handle.net/11245/2.35516

File ID uvapub:35516
Filename cc1.pdf
Version unknown

SOURCE (OR PART OF THE FOLLOWING SOURCE):
Type report
Title Optimal taxation of human capital and credit constraints
Author(s) B. Jacobs
Faculty UvA: Universiteitsbibliotheek
Year 2002

FULL BIBLIOGRAPHIC DETAILS:
http://hdl.handle.net/11245/1.419680

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content licence (like Creative Commons).

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
(pagedate: 2014-11-27)

Optimal Taxation of Human Capital and Credit Constraints*

Bas Jacobs†

Abstract

We study optimal linear income taxation in a model with heterogeneous agents where earnings potentials are endogenously determined through human capital accumulation. Agents differ in initial conditions and ability to learn. Capital market imperfections prevent poor agents to invest optimally in human capital. We show that optimal linear tax rates on human capital are positive, even in absence of redistributive preferences of the government. A more progressive tax system has efficiency gains because credit constraints are relaxed. Numerical calculations show that optimal linear tax rates are significantly increased when capital market imperfections are present.

Key words: optimal linear taxation, human capital, credit constraints
JEL codes: H21, H23, J24

*First draft: February, 1999. This draft: March, 2002. I thank Lans Bovenberg, Casper van Ewijk, Ruud de Mooij, Paul Tang, Peter Sørensen, and Coen Teulings for helpful comments and suggestions. I also benefitted from suggestions made by participants of seminars held at the Tinbergen Institute Amsterdam, The CPB Netherlands Bureau for Economic Policy Analysis, and the 56th conference of the International Institute for Public Finance held in Seville, Spain, August 28-31, 2000. Any remaining errors are mine. Financial support from the NWO Priority Program 'Scholar' financed by the Dutch Organisation of Sciences is gratefully acknowledged.

†University of Amsterdam, Tinbergen Institute, and CPB Netherlands Bureau for Economic Policy Analysis. Address: Department of General Economics, Faculty of Economics and Econometrics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. Phone: (+31) - 20 - 525 5088. Fax: (+31) - 20 - 525 4310. E-mail: jacobs@fee.uva.nl.

1 Introduction

Friedman and Kuznets (1945) were among the first to recognize that human capital is illiquid and makes bad collateral. Therefore, lenders provide funds only to a very limited extent to finance these investments. Investments in human capital may be sub-optimally low as a consequence. Credit markets imperfections can have adverse effects on the distribution of income, economic growth and community formation as some important papers of more recent times suggest.

Intergenerational mobility of human capital is reduced if parents invest sub-optimally in their children when they cannot borrow the funds to finance their children's education, see Becker and Tomes (1979, 1986) and Loury (1981). Not only inequality may be increased but also economic activity can be reduced if capital markets fail. If economic growth depends in one way or another on the level of human capital an economy accumulates, credit market imperfects may reduce the rate of economic growth, see e.g. Galor and Zeira (1993) and Perotti (1993). Capital market market imperfections can also contribute to segregation of communities. Poor households may not be able to leave their (poor) communities as their life-time levels of human capital are too low as a consequence of sub-optimal investments in human capital. Local externalities, taxes, and feedbacks of human capital add to the persistence of income inequality, see e.g. Bénabou (1996a, 1996b), Durlauf (1996), and Fernàndez and Rogerson (1996, 1998).

Empirical evidence on the importance of capital market imperfections is based on the notion that when investments in human capital are constrained, there should be a correlation between parental incomes and education of children, see also Becker and Tomes (1979, 1986). This correlation is found indeed, see for example Bishop (1977), Lazear (1983), Corcoran, Gordon, Laren, and Solon (1992), Kane (1994), and the overview in Haveman and Wolfe (1995). However, the correlation between parental incomes and children's education might not be the consequence of failing of credit markets, but also numerous other factors such as the education of parents. After instrumenting for parental incomes to correct for this bias, Shea (2000) finds, that parental incomes remain to exert an influence on children's education in the bottom quartile of the income distribution. Plug and Vijverberg (2000) also find strong empirical support for the importance of credit constraints

after controlling for the endogeneity of parental incomes.[1]

Some papers suggest that redistributive policies are in general efficient if capital markets fail, e.g. Bénabou (1996a, 1996b), Fernàndez and Rogerson (1996, 1998). The idea is that redistributive polices allow poor agents to invest optimally in human capital. However, the intuition for this result is less well understood. Redistribution generally distorts economic decisions. As stressed by the traditional optimal taxation literature redistribution entails efficiency losses, see e.g. Mirrlees (1971). In this paper we answer the question why it is so that redistribution is beneficial even if it goes along with efficiency losses.

In this paper we analyze optimal income taxation when credit constraints are relevant. Many authors have analyzed the effects of taxation on learning decisions see, for example, Stokey and Rebelo (1995), Milesi-Feretti and Roubini (1998), Jones, Manuelli and Rossi (1993, 1997) and Judd (1999). However, none of these papers paid attention to the consequences of credit constraints.[2] We augment the theory on optimal taxation of human capital using a model in similar vain as the model in Eaton and Rosen (1980).

The first point of this paper is to show that even in the absence of redistributional concerns the optimal tax schedule is progressive when credit constraints are binding. The intuition is that a more progressive tax system redistributes incomes from the old (the rich) to the young (the poor). The young face credit constraints so that these are relaxed by a more redistributive tax system and thereby increases investments in human capital.[3] The latter causes a *first*-order welfare gain as the learning decisions were initially distorted. The tax, on the other hand, initially causes only *second*-order welfare losses, as there was no tax to begin with. Consequently, welfare can be improved as the distortionary tax corrects for the initial non-tax distortion in the capital market.

The second contribution of this paper is to add redistributional concerns to show that the pursuit of equality of incomes yields more equality of op-

[1]This is not uncontroversial, however. Cameron and Heckman (1999) and Cameron and Taber (2000) argue that borrowing constraints are not important.

[2]Nielsen and Sørensen (1997) is one of the few papers to analyze optimal taxation in a model with human capital accumulation and liquidity constraints. Notwithstanding that they reach ambiguous conclusions. The reason is that they study fiscal taxation rather than redistributive taxation.

[3]The linear tax schedule is a progressive one as the average tax rate increases with income.

portunity. The latter is defined as the absence of credit market imperfections so that everyone with sufficient ability is able to invest in human capital at the same conditions. Both low income agents and high ability agents benefit from a progressive tax schedule. Although, high ability agents are the potentially high income earners, they suffer the most from the credit constraints, *ceteris paribus*. Therefore, they reap the highest efficiency gains from a redistributive tax schedule. Bénabou (2000) analyzes the political economy of redistributive policies in a model with human capital accumulation, heterogeneous agents, and credit constraints. He derives that redistributive policies that lead to efficiency gains, that is, alleviation of credit constraints, can receive a majority support. The results presented in this paper can be interpreted as the welfare economics counterpart of Bénabou (2000).

The third contribution is that we provide quantitative evidence for the importance of credit constraints for the setting of the optimal tax schedule. Numerical calculations indicate that optimum income tax rates are significantly higher when capital market imperfections are present compared with optimum tax rates derived under perfect capital markets.

This paper is related to the existing literature in a number of ways. First, Hubbard and Judd (1986) have found a similar result in a model without human capital formation. They numerically analyze the effects of tax policies in a dynamic model where agents are liquidity constrained. They find that taxing income progressively mitigates the adverse welfare effects of credit market imperfections on savings. Second, the optimality of progressive taxation in a second-best world has been found as well in Eaton and Rosen (1980). Uninsurable risks of investments in human capital gives rise to a non-tax distortion in the economy. Risk averse agents under-invest in human capital from a social point of view as they require a risk-premium on their investments in human capital. Progressive income taxation is optimal because agents are partially insured against the uncertain outcomes of investments in human capital. Third, the typical second-best result of the optimality of progressive taxation can also be related to the models of taxation in labor markets with frictions, unions, or efficiency wages, see for example Bovenberg and Van der Ploeg (1994), Pissarides (1998), and Sørensen (1999). A progressive tax schedule in these imperfect labor markets moderates wages and improves employment or welfare.

This paper is organized as follows. Section 2 describes the model and individual behavior, in Section 3 optimal fiscal policy is derived, Section 4 discusses some numerical examples and Section 5 concludes.

2 Model

Consider a two-period life-cycle model of human capital formation. Wage rates per unit of human capital and interest rates are taken as given.[4] A mass of agents with unit measure lives for two periods. Agents are heterogeneous with respect to ability to learn α and an endowment ω that accrues to them at the beginning of their lives. One may also view the endowment as a parameter reflecting not only financial but also as a monetary short-cut for the value of non-liquid assets such as the network of parents, or the presence of collateral such as houses, that may reflect better access to the capital market. Agents with a higher ability to learn are relatively more efficient in the production of human capital. The higher is ω, the more favorable are initial conditions. The cumulative joint distribution of α and ω is denoted by $F(\alpha, \omega)$. F has supports $[\underline{\alpha}, \infty)$ and $[\underline{\omega}, \infty)$.

In the first period agents decide to allocate their time between working and learning, in the second period they work fully and consume their savings (if they have any savings). We abstract from an endogenous labor supply decision.[5] Every agent has one unit of human capital at the beginning of its life. A fraction x of total time in the the first period is spend on education, the rest $1 - x$ is devoted to working, where the total time endowment is normalized at unity. We assume that education requires, besides time, κ market goods per year of education. For simplicity, we do not allow for substitution between time and goods invested in education.[6] Total human capital h is a function ϕ of ability α and time x invested in education:

$$h_{\alpha\omega} = \phi(\alpha; x_{\alpha\omega}), \tag{1}$$

with $\phi_\alpha, \phi_x > 0$, $\phi_{\alpha\alpha}, \phi_{xx} < 0$, and $\phi_{\alpha x} = \phi_{x\alpha} > 0$. The subscript refers to the argument of differentiation. There are positive, but diminishing returns to ability and time invested in education. Agents with higher ability levels

[4] The partial equilibrium model can nevertheless be thought of as the steady state of a small open economy in which perfect capital mobility fixes the real interest rate.

[5] Although the model becomes highly non-linear, we do not expect that qualitative results change when labor supply is endogenous. We expect that effective labor supply elasticities go up however. Jacobs (2001) analyzes optimal income taxation with endogenous human capital and labor supply with perfect capital markets and finds that optimal taxes should be lower since tax distortions are increased by an endogenous labor supply decision.

[6] Qualitative effects are not affected by allowing for substitution between goods and time invested in education.

are assumed to invest relatively more time and goods in human capital accumulation as indicated by the positive cross-derivative. Diminishing returns to investments in human capital ensure an interior solution with a constant interest rate.

Income derived from working equals $(1-t)w(1-x)$, where t is the flat labor income tax rate, and w stands for the wage rate. We restrict the analysis to linear tax rates only in order to reduce the analytical complexities involved. Perfectly equalizing differences in wage rates are assumed so that wages per unit of human capital are equal.

The tax authority is unable to observe α and ω so that individualized lump-sum transfers are excluded. The non-observability of α is standard. The non-observability of ω is not: we assume that the endowment is not subject to the income tax. The analysis is qualitatively not affected by this assumption. Two reasons can be given for making this assumption. First, parental incomes are usually already taxed before the children receive the transfers and we do not model the role of parents explicitly. Second, as ω also reflects non-monetary means, it is hard, if not impossible, for the tax-authority to tax these. Furthermore, the tax authority cannot distinguish between income from raw labor and human capital. The latter is the analogue to the non-observability of wage rates and hours worked. Consequently, taxes on both human capital and labor are both equal to t.

Every agent might receive a non-individualized lump-sum income transfer, or negative income tax, g, in both periods of its life. With positive transfers g the tax-system is progressive since the average tax rate increases with income. Goods invested in education κx are non-tax deductible. Savings s equal total income minus first period consumption c_1 and the direct costs of education κx. The consumption price is chosen as the numéraire. Consumption is not taxed. The first period budget constraint is therefore given by:

$$c_{1\alpha\omega} + \kappa x_{\alpha\omega} + s_{\alpha\omega} = (1-t)w(1-x_{\alpha\omega}) + \omega + g, \tag{2}$$

In the second period, income derived from accumulation of financial assets is $(1+r)s$ where r is the constant real interest rate. Interest income is untaxed.[7] All income from human and financial sources is used for second

[7]Taxes on savings encourage human capital formation because real interest costs are lower, see Heckman (1976). We abstract from the analysis of optimal dual income taxation where taxes on labor income and capital income are simultaneously optimized, see e.g. Nielsen and Sørensen (1997) for a model with homogeneous agents and dual income

period consumption c_2. Hence, the second period budget constraint is:

$$c_{2\alpha\omega} = (1-t)wh_{\alpha\omega} + (1+r)s_{\alpha\omega} + g. \tag{3}$$

We assume that second period consumption is always higher than first period consumption. This is in conformity with reality. The consequence of this assumption is that the lump-sum transfer affects first period consumption more than second period consumption, so that a progressive tax-schedule, entails redistribution from the old (the rich) to the young (the poor).

Agents with low initial incomes are subject to credit market imperfections. The credit market fails because human capital cannot serve as collateral for loans as individuals cannot sell claims on their future incomes to banks - slavery is abandoned, see Friedman (1962). Moreover, the credit market is likely to fail as well because of information asymmetries - causing problems with moral hazard and adverse selection in credit markets. Credit market imperfections can take various forms. For analytical tractability, the common assumption is made that the credit market imperfection takes the form of a non-negativity constraint on financial savings:[8]

$$s_{\alpha\omega} \geq 0. \tag{4}$$

Finally, we do not allow for education subsidies or loans, as in Eaton and Rosen (1980).[9] Of course, most governments give education subsidies and loans on a large scale in most developed countries in order overcome problems in imperfect capital markets. A priori there is no reason to exclude these as an instrument. Nevertheless, we observe that, even in the presence of the highly developed systems of education finance, there is empirical support for the presence of liquidity constraints for the poor agents. We could therefore set the level of subsidies at some prefixed level, that is insufficient to eliminate all credit constraints, but this would not affect our main argument. Additionally, our argument gains relevance in less developed countries where systems of

taxation. Bovenberg and Jacobs (2001) analyze optimal dual income taxation with heterogeneous agents in a comparable model without credit constraints. Eaton and Rosen (1980) proceed in similar fashion.

[8] It requires a complex model to allow explicitly for uncertainty and asymmetric information. See for example Stiglitz and Weiss (1981) for the analysis of credit rationing in markets with imperfect information.

[9] See Bovenberg and Jacobs (2001) for an analysis of the simultaneous setting of optimal taxes and education subsidies in a similar model as the one presented here.

education finance are not that well developed, and we can expect that credit constraints impose large restrictions to agents to invest optimally in human capital.

Moreover, Mincer (1962) and Heckman et al. (1998) suggest that a large part of human capital formation can be attributed to on-the-job-training. Since these training efforts are generally not verifiable, subsidies on on-the-job training are hard to target efficiently due to the moral hazard and adverse selection problems induced by non-verifiable efforts, see also Van Ewijk and Tang (2000). Training subsidies may therefore not be so effective so as to eliminate all credit constraints.

Utility u is a concave twice differentiable utility function with positive but diminishing marginal utility of consumption in both periods:

$$u(c_{1\alpha\omega}, c_{2\alpha\omega}), \tag{5}$$

and $u_{c_1}, u_{c_2} > 0$, $u_{c_1c_1}, u_{c_2c_2} < 0$, $u_{c_1c_2} = u_{c_2c_1} \geq 0$.

Agents maximize utility by choosing consumption c_1, c_2, and the optimal amount of learning x, subject to their budget constraints, the production function of human capital and the non-negativity constraint on savings. Since the optimization problem is well-behaved, first order conditions are both necessary and sufficient. Manipulation of the first order conditions gives the Frisch demand equations (see also the appendix):

$$c_{1\alpha\omega} = c_1\left[\lambda_{1\alpha\omega},\ \lambda_{2\alpha\omega}\right], \tag{6}$$

$$c_{2\alpha\omega} = c_2\left[\lambda_{1\alpha\omega},\ \lambda_{2\alpha\omega}\right], \tag{7}$$

$$\begin{aligned} \lambda_{1\alpha\omega} &= (1+r)\lambda_{2\alpha\omega}, \quad s_{\alpha\omega} \geq 0, \\ \lambda_{1\alpha\omega} + \mu_{\alpha\omega} &\geq (1+r)\lambda_{2\alpha\omega}, \quad s_{\alpha\omega} < 0, \end{aligned} \tag{8}$$

with $c_{1,\lambda_1}, c_{2,\lambda_2} < 0$, $c_{1,\lambda_2} = c_{2,\lambda_1} \geq 0$, where the subscript refers to the argument of differentiation. λ_1 and λ_2 stand for marginal utility of income in periods 1 and 2 respectively. μ is the Lagrange multiplier associated with the non-negativity constraint on savings. It is positive if the constraint is binding, and equal to zero if the constraint is slack. In the last case, the price of period 1 consumption relative to period 2 consumption is simply equal to $1+r$. If the non-negativity constraint is binding, the marginal value of income in period one increases, so that the price of consumption in period one relative to period two satisfies $\lambda_1 + \mu \geq (1+r)\lambda_2$.

Conditional demand for investment inputs x in the production of human capital is written as:

$$x_{\alpha\omega} = x\left[\frac{\lambda_{1\alpha\omega}((1-t)w+\kappa)}{\lambda_{2\alpha\omega}(1-t)w};\alpha\right], \tag{9}$$

with $\frac{\partial x}{\partial R} < 0$, $R \equiv \frac{\lambda_{1\alpha\omega}((1-t)w+\kappa)}{\lambda_{2\alpha\omega}(1-t)w}$, since $\phi'' < 0$. By virtue of the concavity of the production function of human capital and the complementarity between ability and x in the production of human capital we may sign $\partial x/\partial\alpha > 0$. More talented agents learn more as their productivity of time invested in human capital accumulation is higher ($\phi_{x\alpha} > 0$).

If the non-negativity constraint is slack, i.e. when $\lambda_1/\lambda_2 = 1 + r$, the optimal choice of inputs is not affected by initial conditions as a result of the separation between optimal investment and consumption decisions. Consequently, we may sign the partial $\partial x/\partial\omega = 0$ if $s \geq 0$. Agents with an initial endowment $\omega \geq \omega^o \equiv c^u_{1\alpha\omega} + \kappa x^u_{\alpha\omega} - (1-t)w(1-x^u_{\alpha\omega}) - g$ are not credit constrained in the first period. c^u, and x^u denote optimal first period consumption and time devoted to human capital accumulation if savings are positive. Time invested in human capital is thus equal for every agent with the same ability α if agents are unconstrained, although ω may differ.

Since binding credit constraints prevent agents to invest optimally in human capital, we can establish that $\partial x/\partial\omega > 0$ if $s < 0$. Agents have to sacrifice future consumption for their investments in human capital. This becomes increasingly more expensive in terms of utility as a consequence of diminishing marginal utility. As such, the lower are initial incomes, the higher is the relative price of future income, and the lower are investments in human capital. From the definition of ω^o is clear that, either if agents are very poor (low ω) or have high ability (high α) and learn a lot, credit constraints become more important.

Note further that taxes distort investments in human capital. A higher tax rate decreases investments in human capital since total costs - foregone earnings and goods invested - are less affected by changes in the tax rate, than the returns, in the form of future earnings. If goods were not an input in production of human capital, or were fully tax deductible, the tax rate would have no effect on investments in human capital since costs and returns are equally affected by the tax.[10]

[10] Alternatively, we could introduce a (similar) tax distortion on learning decisions by

3 Optimal taxation

3.1 Optimal taxation without redistribution

First, we show that optimal linear tax rates are positive even in the absence of redistributive preferences *and* the use of lump-sum transfers. However, these transfers are not age-dependent which is crucial for our exposition. To that end we consider a model with a representative agent. There are two instruments at the disposal of the government: the linear tax rate on labor income t and the negative income tax g. The government collects taxes to finance the negative income tax and a revenue requirement Λ.[11] We assume that the government can indeed borrow funds at real interest rate r to satisfy the inter-temporal budget constraint. The reason is that the government has a claim on all human capital stocks through the tax system as it is able to enforce tax payments. Therefore it can circumvent the capital market constraint that individuals face.

The government budget constraint reads as:

$$\Lambda + G = tH, \tag{10}$$

where $G \equiv g(2+r)$, and $H(x) \equiv w(1-x) + w\phi(x)/(1+r)$ stands for the lifetime value of human capital discounted at rate r. It's analytically convenient to work with $H(x)$ as a measure for human capital (or labor income).

Note that $\partial H/\partial G = 0$ if the agent is not constrained, since the investment decision is not affected by taxes. $\partial H/\partial G > 0$ if the agent is liquidity constrained, since a relative increase in first period income relaxes the credit constraint, so that learning time increases. Due to the progressive tax schedule, the government allows the agent to transfer resources from second period to the first period of its life. Moreover, the lump-sum element has distortionary effects as the relative price of current in terms of future consumption (λ_1/λ_2) is affected by G.

employing a tax schedule where marginal rates on tax on second period income are higher than marginal rates on first period income as in Nielsen and Sørensen (1997). We could also allow for endogenous labor supply in the second period so as to get a tax distortion on learning through the utilization of human capital as in Eaton and Rosen (1980). If taxation induces agents to work less, returns on investments human fall as the utilization rate of human capital decreases. Both modifications complicate the analysis without altering the main conclusions of this paper.

[11] Revenue requirements are not necessary for our exposition, but we follow common practice here.

Further, $\partial H/\partial t < 0$.[12] The tax rates t and lump-sum transfers G are chosen so as to maximize the indirect utility function of the representative agent $V(t, G) \equiv u(c_1^*, c_2^*)$, subject to the first order condition for optimal learning and the government budget constraint. The Lagrangian is:

$$\mathcal{L} = V + \eta(tH - G - \Lambda), \tag{11}$$

where η is the Lagrange multiplier associated with the government budget constraint. The first-order condition for the lump-sum element G is:

$$\frac{\partial \mathcal{L}}{\partial G} = \lambda_1 - \eta + \eta t \frac{\partial H}{\partial G} = 0, \tag{12}$$

where Roy's lemma is used: $\partial V/\partial G = \lambda_1$. From this follows that the marginal costs of public funds ($MCPF$), defined as η/λ_1, are larger than unity since $\frac{\partial H}{\partial G} > 0$:

$$MCPF \equiv \frac{\eta}{\lambda_1} = \frac{1}{1 - t\frac{\partial H}{\partial G}} > 1. \tag{13}$$

If there were no credit constraints, we have $\frac{\partial H}{\partial G} = 0$, and the $MCPF$ equals unity. The marginal costs of funds increase if human capital reacts more sensitively to an increase in the lump-sum taxation, i.e. when $\frac{\partial H}{\partial G}$ increases. The reason is that credit constraints become more tight when the government employs lump-sum element to generate revenues. The last expression reveals that lump-sum finance is distortionary.

The first-order condition for the tax rate is given by:

$$\frac{\partial \mathcal{L}}{\partial t} = -\lambda_1 w(1 - x) - \lambda_2 w\phi(x) + \eta H + \eta t \frac{\partial H}{\partial t} = 0, \tag{14}$$

where Roy's lemma gives: $\partial V/\partial t = -\lambda_1 w(1 - x) - \lambda_2 w\phi(x)$. We rewrite the last expression:

[12] For low t, $\partial H/\partial t$ may be positive for liquidity constrained agents, because there are two opposing effects. First, there is the negative effect of taxation due to the tax distortion on learning. Second, taxing income affects second period income relatively more than first period income. Therefore, λ_1/λ_2 decreases and learning time increases. At the optimum, however, we must have that the negative effect on learning dominates the positive effect and $\partial H/\partial t < 0$. Otherwise, welfare could be increased by increasing t which is inconsistent with the notion of an optimum. See the appendix for the derivations of the comparative statics.

$$-\lambda_1 w(1-x) - \lambda_2 w\phi(x) = -\lambda_1 H(1-q), \tag{15}$$

where $q \equiv \left(\frac{1}{1+r} - \frac{\lambda_2}{\lambda_1}\right)\chi > 0$, and $\chi \equiv w\phi(x)/H$ is the share of second period income in total human capital. q denotes the welfare loss due to the presence of the liquidity constraint. q is strictly non-negative, since with liquidity constraints we have $\lambda_2/\lambda_1 < 1/(1+r)$. Using the last result in the first-order condition gives the optimal linear tax rate:

$$\frac{t}{1-t} = \frac{1}{\varepsilon}\left(1 - \frac{1}{MCPF} + \frac{q}{MCPF}\right), \tag{16}$$

where $\varepsilon \equiv -\frac{\partial H}{\partial t}\frac{1-t}{H} > 0$ is the uncompensated elasticity of life-time earnings with respect to the tax rate. This expression has an intuitive interpretation. First, the larger the elasticity of the labor income tax base with respect to the tax rate, the lower should the optimal tax be, because taxation causes efficiency losses.

Second, the first two terms in the brackets measure the costs of employing the lump-sum element in revenue raising. $1 - 1/MCPF$ is positive due to the fact that the $MCPF$ are larger than unity. The higher are the $MCPF$, the more costly it is to raise revenues with the lump-sum element as credit constraints become more tight. Therefore, the government relies more on the tax rate to finance revenues accordingly.

Third, the last term measures the extent to which the credit constraint calls for redistributive taxation. If credit constraints are more tight, q increases, so that the optimal tax rate on labor income increases. By increasing the tax rate the government allows the agent to transfer future income to the first period of its life.

The intuition for the fact that a progressive tax system is always optimal if there are credit constraints is as follows. Since there is a non-tax distortion (credit constraint) in the economy in the absence of taxation, it is always optimal to reduce this distortion through G with the distortionary tax, because the latter initially only causes second order welfare losses, since there was initially no tax, whereas the relaxation of the credit constraint causes a first-order welfare gain since the learning decision was already distorted as a consequence of the credit constraint.

If there are no credit constraints, the lump-sum element can be employed in non-distortionary fashion and $MCPF$ equals unity. q is zero as well in

that case. Hence, in the absence of credit constraints, the optimal tax rate on labor income is zero

Proposition 1 *In the absence of redistributive motives, the optimal tax rate is strictly positive if credit constraints prevent agents to invest optimally in human capital. The optimal tax rate t should be higher if: i) the $MCPF$ are higher, that is, when the transfer G is less suited to raise revenues as this instrument aggravates credit constraints; ii) credit constraints are more important, i.e. q is larger; iii) the uncompensated elasticity of human capital ε w.r.t. the tax rate is lower. If credit constraints are not binding, the $MCPF = 1$, and $q = 0$, so that $t = 0$.*

3.2 Optimal taxation with redistribution

In this section we turn to the more general case of optimal income taxation with heterogeneous agents so that redistributive motives of the government explicitly enter through the social welfare function. The arguments derived in the previous section can be applied in straightforward fashion to this case. The appendix contains the derivation. The government budget constraint now reads as:

$$\Lambda + G = t \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} H_{\alpha\omega} dF(\alpha, \omega), \tag{17}$$

where $H(x_{\alpha\omega}) \equiv w(1 - x_{\alpha\omega}) + w\phi(\alpha; x_{\alpha\omega})/(1 + r)$.

The tax rate t and lump-sum transfers G are chosen so as to maximize a social welfare function Γ over individual utilities:

$$\Gamma = \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi(V_{\alpha\omega}) dF(\alpha, \omega), \quad \Psi' > 0, \quad \Psi'' \leq 0, \tag{18}$$

where $V_{\alpha\omega}$ is the indirect utility function of the agents, subject to the first order condition for optimal learning and the government budget constraint. Different assumptions about Ψ yield e.g. a Rawlsian objective function or an utilitarian objective function ($\Psi' = 1$), see also Atkinson and Stiglitz (1980).

From, the optimum condition for G we can derive the marginal costs of public funds, which is now given by η/λ, where $\lambda \equiv \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 dF$ is the average of the marginal social valuation of income:

$$MCPF \equiv \frac{\eta}{\lambda} = \frac{1}{1 - \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} t \frac{\partial H}{\partial G} dF} > 1. \tag{19}$$

The interpretation of $MCPF$ is the same as before. The lump-sum element has costs in terms of revenue raising because credit constrained agents learn less as G is increased. Therefore, G is distortionary finance.

To find an implicit expression for the optimal tax rate we define the weighted average uncompensated elasticity of life-time earnings with respect to the tax rate:

$$\bar{\varepsilon} \equiv \frac{\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \varepsilon_{\alpha\omega} H_{\alpha\omega} dF}{\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} H_{\alpha\omega} dF} > 0, \tag{20}$$

where the weights are the income levels of the agents. Elasticities of agents with higher income levels are weighted more because the average distortion of taxation increases if agents have higher incomes.

Further, we introduce the distributional characteristic ξ that comprises the distributional impact that human capital has on social welfare, see also Atkinson and Stiglitz (1976):

$$\xi \equiv -\left(\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \left(\frac{H_{\alpha\omega}}{\bar{H}} \right) \left(\frac{\Psi' \lambda_{1\alpha\omega})}{\lambda} \right) dF(\alpha, \omega) - 1 \right) > 0. \tag{21}$$

The term in brackets measures the normalized covariance of human capital and the marginal social valuation of income. $\bar{H} \equiv \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} H_{\alpha\omega} dF(\alpha, \omega)$ stands for the average human capital, and λ is the average social marginal value of income. The distributional characteristic is positive: $\xi > 0$. The reason is that $\Psi' \lambda_{1\alpha\omega}$ decreases if human capital earnings $H_{\alpha\omega}$ increase as a consequence of diminishing marginal utility of income. i.e. it becomes socially less efficient to redistribute resources to richer agents, given that the social welfare function features diminishing marginal social welfare in utility of the agents. Consequently, the term in brackets is negative.

If the government was not interested in redistribution, every agent has the same marginal social value of income: $\Psi' \lambda_1 = \lambda$, which yields the result that $\xi = 0$ in that case. The distributional characteristic increases if incomes become more unevenly distributed, or if a greater weight is given to agents at the lower end of the income distribution.

The optimal tax rate follows after manipulation of the first order condition for t:

$$\frac{t}{1-t} = \frac{1}{\bar{\varepsilon}} \left(\xi + \frac{(MCPF - 1)}{\overline{MCPF}} + \overline{\left(\frac{q}{MCPF} \right)} \right), \tag{22}$$

where $1/\overline{MCPF} \equiv \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 H dF / \eta \bar{H} > 0$ is the income weighted inverse of the $MCPF$. This term weights the welfare losses from employing

the lump-sum element with the levels of income. The more high income earners, i.e. the agents with the high abilities, react to decreasing the lump-sum element, the more the average inverse of $MCPF$ falls, and *vice versa*. $\overline{\left(\frac{q}{MCPF}\right)} \equiv \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 q H dF / \eta \bar{H} > 0$ is the income weighted social welfare loss associated with the capital market imperfection. The more the high income earners suffer from the credit constraints, the higher the average marginal social welfare loss is above the non-weighted average welfare loss.

Also this optimum tax formula has an intuitive interpretation. Again, the optimal tax rate on labor income should be lower if the average compensated elasticity of human capital with respect to the wage rate is higher as indicated by $\bar{\varepsilon}$.

The first term, ξ, measures the extent to which the government wants to redistribute incomes. The larger is the spread in incomes, and the more concave is the social welfare function, the larger is the distributional characteristic. Hence, the government sets a higher optimum tax if ξ is larger. The formula clearly shows the trade-off between equity (ξ) and efficiency (ε) considerations.

The second term measures the distortionary aspects of using the lump-sum element in raising revenues. The more distortionary lump-sum taxes are, i.e. when the $MCPF$ are higher above 1, the higher the government should set the tax rate, in order to off-set the distortions associated with the lump-sum element. Since the high ability agents suffer relatively more from credit constraints, for given ω, because they learn more, the average inverse of the $MCPF$ falls relatively more than the $MCPF$ increases. This makes that the distortionary impact of employing the lump-sum element in revenue raising is increased. In other words, if more high-ability agents are credit constrained, everything else equal, the government uses less lump-sum finance and the optimum tax should be higher.

The third term measures the welfare costs of the capital market imperfections. The higher is the average welfare loss, the higher the government should set the tax. A more progressive tax therefore alleviates credit constraints. Again, the marginal welfare loss, is aggravated if the relatively high income earners (due to high ability) are also the ones suffering from credit constraints. This makes that income weighted loss of credit constraints increases more than the non-weighted welfare loss as credit constraints are more severe.

If there is a strong correlation between initial income endowments and

ability, which is the empirically plausible case, the distortionary effects of credit constraints are lower. The reason is that *ceteris paribus* high ability agents suffer the most from credit constraints. If these high ability agents are also better endowed, the role of the capital market imperfections is reduced. This results both in a decline in average $MCPF$ and a decline in the average welfare loss due credit constraints. the last two terms in the tax formula. Consequently, optimal tax rates could be lower if the correlation between ability and endowments increases. However, a stronger correlation between initial incomes and ability increases the distributional characteristic, so that optimal tax rates increase on that account. If there were no distortions due to capital market imperfections, a higher correlation between endowments and abilities would of course result in higher optimum taxes as only the distributional characteristic is affected.

If there is no income inequality (H equal for all agents), the government places the same value of income on every agent, $\xi = 0$. In that case, the welfare weighted inverse of the $MCPF$ coincides with the inverse of $MCPF$, and the welfare weighted loss of the capital market imperfection equals $q/MCPF$ for all agents. In this case we have the formula that we obtained earlier in the case of a representative agent, see equation (16).

If the capital market imperfection is absent, there are no income effects of lump-sum transfers, so that $MCPF = 1$.[13] Moreover, in that case, $q = 0$ so that $\overline{\left(\frac{q}{MCPF}\right)} = 0$ and the optimal tax rate equals:

$$\frac{t}{1-t} = \frac{\xi}{\bar{\varepsilon}}. \tag{23}$$

This is a familiar looking expression in linear taxation models showing the trade-off between equity and efficiency, see e.g. Dixit and Sandmo (1977), Atkinson and Stigitz (1980), or Tuomala (1985). The numerator gives the normalized covariance of the marginal utility of income and earnings and the denominator gives the average compensated elasticity. In contrast with the case of a representative agent, the optimum tax remains to be positive even if capital markets are perfect. The reason is that there is still income inequality

[13]It might seem surprising that in the last case the $MCPF$ are unity even if taxation is distortionary. There is, however, nothing intrinsically special to this case. In Atkinson's (1995) analysis where optimal linear taxation of income is analyzed in a model with endogenous labor supply, and preferences are such that income effects are absent - as in our analysis in the absence of credit constraints -, the $MCPF$ are unity as well.

due to differences in ability, so that the government wants to redistribute incomes.

Proposition 2 *In the presence of redistributive motives, the optimal tax rate is strictly positive, whether credit constraints prevent agents so invest optimally in human capital or not. The optimal tax rate t should be higher if: i) the $MCPF$ are higher, that is, when the transfer G is less suited to raise revenues as this instrument aggravates credit constraints; ii) credit constraints are more severe, i.e. q is large; iii) the average uncompensated elasticity $\bar{\varepsilon}$ of human capital w.r.t. the tax rate is lower; iv) if the government attaches greater weight to income equality, so that ξ is higher. If credit constraints are not binding for all agents we have that $MCPF = 1$, and $q = 0$, so that the tax rate t is determined only by the distributional characteristic ξ and the elasticity $\bar{\varepsilon}$.*

4 Numerical examples

This section considers some numerical examples of the optimal tax rates. The main focus in our calculations is on effects of credit constraints on optimum tax rates. The method employed here stems from Stern (1976). At forehand must be noted that these results should be interpreted with caution since there is a large number of (unknown) parameters involved.

We make two simplifying assumptions so as to solve the model analytically for the consumers in the cases where capital market imperfections are either present or absent, see the appendix.[14] First, we assume that the tax code is such that only in the first period a lump-sum transfer is given. This assumption can be justified if the share of the second period transfer in total second period income is low, which is not unreasonable. Second, the utility function is logarithmic, i.e. with an elasticity of substitution equal to unity:

$$u(c_1, c_2) = \ln c_1 + \ln c_2, \tag{24}$$

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

$$\phi(\alpha, x) = A\alpha^{\beta} x^{\gamma}, \tag{25}$$

[14] This is done for computational convenience because the model may very easily become very complex to solve numerically.

where $A > 0$ is the exogenous productivity of learning. The Cobb-Douglas function is used in almost the entire human capital literature, see e.g. Weiss (1986) or Trostel (1993). β is the elasticity of ability in learning and is calibrated to give a realistic spread in learning time.

The social welfare function is a Samuelson-Bergson utility function with a constant elasticity of inequality aversion υ:

$$\Gamma = \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \frac{V^{1-\upsilon} - 1}{1 - \upsilon} dF(\alpha, \omega). \tag{26}$$

If $\upsilon = 0$ the social welfare function is utilitarian, if $\upsilon = \infty$ the social welfare function is Rawlsian, see also Atkinson and Stiglitz (1980).

The joint density of ability and endowment incomes is assumed to be bivariate log-normal with means μ_α, μ_ω, standard deviations σ_α, σ_ω, and correlation $\rho_{\alpha\omega}$:

$$(\ln \alpha, \ln \omega) \sim N[\mu_\alpha, \mu_\omega, \sigma_\alpha, \sigma_\omega, \rho_{\alpha\omega}]. \tag{27}$$

The reason for assuming this distribution is that the income distribution is approximately log-normal. In order to generate a log-normal income distribution, abilities must be log-normally distributed.[15] Initial endowments are also likely to be drawn from a log-normal income distribution since parental incomes are also approximately log-normally distributed. Moreover, the joint normal distribution allows us to construct an artificial data-set in an easy way: for each pair (α,ω) we can compute the density $f(\alpha, \omega)$ by conditioning ω on α.[16]

Each period is thought of lasting 30 years. The interest rate is set at $r = 5$ which corresponds to approximately 5.4% per year. Wage rates per unit of human capital w are harmlessly normalized at unity. Taxes are solely redistributive because the government revenue requirement is set at $\Lambda = 0$. The pure rate of time preference ρ is set at zero. The social welfare function is utilitarian, so that $\upsilon = 0$.

Direct costs of education are set at $\kappa = .5$ which is half of the wage rate. Therefore, direct costs make up for 1/3 of total costs per year of education.

[15] Alternatively, one may draw ability from a normal distribution and write the production function for human capital as: $\phi(\alpha, x) = A \exp[\alpha]^\beta x^\gamma$.

[16] We constructed an artificial data-set based on 10x10=100 observations representing the deciles for ability and initial income endowments.

Boskin (1975) and Becker (1993) argue that direct costs make up for 1/4 of total costs of education. We use a somewhat higher value here.[17]

The elasticity of time in learning is set at $\gamma = .4$. After reviewing the literature extensively, Trostel (1993) sets this elasticity at .45. The elasticity of ability is set at $\beta = .5$. If the elasticity of ability in the production function for human capital is higher, investments in human capital rapidly diverge between ability groups; varying other parameters often give corner solutions which we want to avoid.

The other parameter for the the human capital technology (A) is calibrated so that the lowest 10 percentile ability agents spend 20% (6 years) in the first period to learning and the highest 10 percentile ability agents spend 80% (24 years) of their time to learning. This results in $A = 17.5$.

The distribution parameters are set as follows. Following e.g. Mirrlees (1971) and Stern (1976), and Tuomala (1990), the mean of the logarithm of ability μ_α is equal to -1 which is a normalization. The standard deviation of ability σ_α is assumed to follow the standard deviation of the income distribution which is roughly 0.4. The same standard deviation is assumed for the initial endowments. The mean of the initial distribution is calibrated so that 26% of the agents are constrained in the absence of taxes and transfers. Shea (2000) finds that credit constraints are relevant for at least the 25% poorest agents. The correlation coefficient between endowments and ability is set at .4. Kremer (1997) and Shea (2000) find that the correlation between children's educational attainment and parental educational attainment is approximately .4. If parental education is a good proxy for initial log endowments a correlation of .4 seems reasonable.

Optimal tax rates under perfect capital markets t^u and under credit market imperfections t^c are derived separately. The reason is that by doing so one can judge the relative importance of credit market imperfections and the implications for optimal policies. Table 1 shows optimal tax rates for different distributional assumptions when the elasticity of the production function γ is varied. Tax rates are derived for different means of initial income endowments. The middle column gives the base-case. In the left column, initial income endowments are lowered, whereas in the right column, they are in-

[17] A reason to justify this value is that we ignored an endogenous labor supply decision that might affect incentives to learn. If taxation reduces labor supply, returns on investments in human capital fall, so that taxation reduces the utilization rate of human capital. By increasing the value of the direct cost parameter we potentially correct for a too low impact of taxation on learning incentives due to ignored labor supply effects.

Table 1: Optimal tax rates (%)

	$\mu_\omega = -2.5$		$\mu_\omega = -1.25$		$\mu_\omega = 0$	
γ	t^u	t^c	t^u	t^c	t^u	t^c
0.1	39.4	39.4	40.5	40.5	43.0	43.0
0.2	29.4	34.3	30.3	31.0	35.2	32.2
0.3	23.7	34.1	24.4	29.5	25.8	25.8
0.4	19.8	33.8	20.4	29.1	21.4	22.4
0.5	17.1	35.0	17.5	30.7	18.2	21.5
0.6	15.2	35.2	15.4	31.7	15.7	22.9

creased.

In the base case scenario where $\gamma = .4$ and $\mu_\omega = -1.25$, the optimal tax rate equals 20.4%. Given the assumptions used, something similar as Stern (1976) found in the case of endogenous labor supply was to be expected. Stern finds an optimal tax rate on labor income of 22.3% when the elasticity of substitution between consumption and leisure equals .4. Tax rates increase if the elasticity of learning time in production of human capital decreases. The reason is that taxes become less distortionary. Similar reasoning holds for increasing the elasticity. With elasticities higher than .6, highest ability agents get into corner solutions as they invest their maximum time endowment in learning.

If the capital market is perfect, optimal tax rates are lower if mean endowments are decreased. The reason is that there is less inequality. The reverse holds for an increase in initial endowments. The changes in optimum tax rates due to changes in endowments are small however. The reason is twofold. Income endowments are not taxed, and endowments only constitute a small fraction of life-time incomes.

If credit markets are imperfect, optimal tax rates are significantly higher than under perfect capital markets. For the base-case the optimal tax rate equals 29.1% which is almost half as high as with perfect capital markets. If initial income endowments are lowered, the tax rate increases to 33.8%, whereas if initial incomes are increased, the optimal tax rate almost equals the tax rate under perfect capital markets (22.4%). The last finding implies that for mean income endowments that are high enough, credit constraints become non-binding.

Apparently, costs of imperfectly functioning capital markets are signifi-

Table 2: Optimal tax rates (%)

	$\mu_\alpha = -2$		$\mu_\alpha = -1$		$\mu_\alpha = 0$	
γ	t^u	t^c	t^u	t^c	t^u	t^c
0.1	38.4	38.4	40.5	40.5	42.3	43.2
0.2	27.5	27.5	30.3	31.0	32.7	42.0
0.3	21.0	21.1	24.4	29.5	27.3	44.0
0.4	16.5	16.5	20.4	29.1	23.9	46.0
0.5	13.0	13.0	17.5	30.7	21.8	49.2
0.6	10.2	10.2	15.4	31.7	20.6	52.3

cant in terms of welfare compared to the welfare costs of taxation. This result is reinforced when the elasticity of learning increases. The optimal tax rates under perfect capital markets decrease because costs of taxation increase. Under imperfect capital markets, the tax rate increases with the elasticity of the learning function. The reason is that efficiency gains of taxation increase as the marginal returns on human capital increase when the elasticity is higher. Costs of taxation therefore increase less than the benefits of taxation as the elasticity rises. Note also that the tax rate varies non-monotonically with γ. A finding that is due to the fact that H is non-monotonically varying with γ.[18]

Mean ability is changed in table 2. If mean ability is lowered, optimal taxes under perfect capital markets are decreased from 20.4 to 16.5% ($\gamma = .4$) because the distribution of income becomes less unequal. The reverse holds if ability increased: taxes increase from 20.4 to 23.9%. Changes in optimal tax rates are relatively small when mean ability changes. Note, however, that increases in ability are accompanied by significant increases in optimal tax rates when there are credit constraints: from 29.1 to 46.0%. Increasing ability make more agents suffer from credit constraints because they learn more. Consequently, efficiency gains from redistributive taxation are high. When ability is decreased the optimal tax rate is sufficient to remove all credit constraints, so that it is equal to the tax rate under perfect capital markets (16.5%).

The first column of table 3 shows optimal tax rates if the elasticity of ability in the production function for human capital β is changed. Under

[18] Stern (1976) found a similar non-monotonic pattern when varying the elasticity of substitution between leisure and consumption in a model with endogenous labor supply.

perfect capital markets the tax rate tends to zero. The reason is that income inequality vanishes - see also proposition 1. An interesting feature is that the optimal tax rate under imperfect capital markets remains largely constant for a large range of parameter values. The optimal tax rate first declines and then increases at high levels of the of the ability elasticity.

The second column gives the optimal tax rates when fixed costs of learning κ are changed. When direct costs are lowered, costs of taxation fall, because taxation becomes less distortionary. Consequently, optimal tax rates increase. This holds for both the tax rates under perfect and imperfect credit markets. The optimal tax rate is almost 50% under perfect capital markets when costs of education are small ($\kappa = 0.1$).

In the third column the interest rate increases from 4 to 12% per year. Higher interest rates reduce optimal investments in human capital. Hence, credit constraints become less important. When the interest rate is 7% or higher, tax rates under perfect and imperfect capital markets coincide as credit constraints cease to be important.

The correlation between initial incomes and ability is varied in the fourth column. Under perfect capital markets, optimal tax rates vary only little with the correlation: there is only a difference of about 5 percentage points in the tax rate when the correlation is increased from .1 to .9. One may conclude that, although there might be a lot of skewness in initial conditions – those with high talents are likely to have good backgrounds – the tax system cannot alter this skewness via the taxes on human capital. The reason is again twofold. First, most variation in the distribution of income is due to education choices. Second, initial endowments are non-taxed. Optimal tax rates remain almost constant as well under imperfect capital markets. The optimal tax rates decline when correlations between ability and endowments increase. In this case the distortionary costs of taxation increase because credit constraints become less important. The reason is that increasing the correlation relaxes credit constraints especially for the high ability agents, who suffer most from capital market imperfections. Hence, the scope for progressive taxes to correct capital market failures decreases. This increase in the distortionary costs of taxation is apparently not off-set by increases social benefits of taxation since there is more inequality with larger correlations between endowments and ability.

In table 4 we have calculated optimal tax rates on human capital where the government requires some revenues Λ. Every .18 point of revenue corresponds to a 10% of total human capital income in the base case scenario. As

Table 3: Optimal tax rates (%)

β	t^u	t^c	κ	t^u	t^c	r	t^u	t^c	$\rho_{\alpha\omega}$	t^u	t^c
.1	2.8	33.8	.1	49.2	49.7	3.32	22.8	41.1	.1	18.6	30.0
.2	7.5	31.0	.2	35.9	42.2	4.48	21.1	32.7	.2	19.2	29.7
.3	12.3	29.4	.3	28.4	36.8	6.05	19.1	23.5	.3	19.8	29.6
.4	16.7	28.9	.4	23.7	32.5	8.17	17.1	17.1	.4	20.4	29.1
.5	20.4	29.1	.5	20.4	29.1	11.02	15.2	15.2	.5	21.0	28.8
.6	23.5	30.5	.6	17.9	26.5	14.88	13.5	13.5	.6	21.6	28.2
.7	26.2	31.3	.7	16.1	24.1	20.09	12.2	12.2	.7	22.2	27.6
.8	28.4	32.4	.8	14.7	22.3	27.11	11.3	11.3	.8	23.0	26.9
.9	30.4	32.9	.9	13.5	20.6	36.60	10.6	10.6	.9	23.8	26.9

Table 4: Optimal tax rates (%)

Λ	t^u	t^c	υ	t^u	t^c
.18 (10%)	21.4	32.5	0	20.4	29.1
.36 (20%)	22.5	36.0	.99	28.9	32.9
.54 (30%)	23.8	40.0	2	34.2	35.9
.72 (40%)	25.3	44.8	3	37.8	38.2
.90 (50%)	27.1	49.6	∞	59.1	59.1

expected, optimal tax rates increase with Λ. However, with imperfect capital markets, increases in tax rates are much larger than with perfect capital markets. The reason is that the poll subsidy is less suited to raise revenues under imperfect capital markets. Therefore, the government relies more on the distortionary tax to raise revenues if capital markets are imperfect.

Preferences of the social welfare function are changed in the last column of table 4. The coefficient of relative inequality aversion υ is increased. Interestingly, at high enough levels of inequality aversion, i.e. with 'Rawlsian' social preferences, optimal tax rates under perfect and imperfect capital markets are equal. The reason is of course that a greater weight on equity induces a more progressive tax-system, but alleviates credit constraints as well. At high levels of inequality aversion no one is rationed and investments in human capital are efficient.

A final remark with regard to the optimum tax rates under credit constraints can be made. Inspection of tables 2, 3, and 4 reveals that optimum tax rates are quite similar - around 30% - when the elasticity of ability β, the

costs of education κ, and the correlation between ability and endowments $\rho_{\alpha\omega}$ are changed. Moreover, even if there is almost no inequality due ability differences, i.e. when β goes to zero, optimum tax rates are still around 30%. These findings echo the Diamond and Mirrlees (1971) result that in an optimal tax-system there should be production efficiency – given the availability of sufficient instruments. In the current, second-best, setup instruments are not sufficient to achieve production efficiency. Nevertheless, it seems that the government corrects for the capital markets imperfection as much as possible. Apparently, this is the case when optimum taxes are around 30%.

5 Conclusion

This paper augmented the theory of optimal income taxation by endogenizing earnings through investments in human capital. Special attention was devoted to the presence of credit market imperfections that prevent (poor) agents to invest optimally. It is shown analytically that the optimal tax schedule is progressive even if redistributional concerns are absent. The reason is that by means of redistribution of resources both from rich to poor and from old to young, credit market imperfections are alleviated. By allowing for redistributional motives, the case for a progressive tax-schedule is strengthened. Notably, both agents with low and high ability gain from redistribution. Low ability agents experience an equity gain. High ability agents suffer more from imperfections on the credit market than low ability agents because they learn more. They consequently experience efficiency gains as their credit constraints are loosened by means of redistribution. These results are confirmed in numerical calculations of optimal tax rates on human capital. The most important finding is that optimal taxes on human capital are significantly increased when credit market imperfections are present.

References

Atkinson, Anthony B. (1995). *Public Economics in Action. The Basic Income/Flat Tax Proposal*, Oxford: Oxford University Press.

Atkinson, Anthony B., and Joseph E. Stiglitz (1976). "The Design of Tax Structure: Direct versus Indirect Taxation", *Journal of Public Economics*, 6, 55-75.

Atkinson, Anthony B., and Joseph E. Stiglitz (1980). *Lectures on Public Economics*, New York: McGraw-Hill.

Becker, Gary S. (1993). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, Third edition, Chicago: Chicago University Press.

Becker, Gary S., and Nigel Tomes (1979). "An Equilibrium Theory of the Distribution of Income and Intergenerational Income Mobility", *Journal of Political Economy*, 87, (6), 1153-1189.

Becker, Gary S., and Nigel Tomes (1986). "Human Capital and the Rise and Fall of Families", *Journal of Labor Economics*, 4, S1-S39.

Bénabou, Roland (1996a). "Equity and Efficiency in Human Capital Investment: The Local Connection", *Review of Economic Studies*, 63, 237-364.

Bénabou, Roland (1996b). "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance", *American Economic Review*, 86, (3), 584-609.

Bénabou, Roland (2000). "Unequal Societies: Income Distribution and the Social Contract", *American Economic Review*, 90, (1), 96-129.

Bishop, John (1977). "The Effect of Public Policies on the Demand for Higher Education", *Journal of Human Resources*, 12, 285-307.

Boskin, Michael (1975). "Notes on the Tax Treatment of Human Capital", NBER Workingpaper 116.

Bovenberg, A. Lans, and Bas Jacobs (2001). "Redistribution and Education Subsidies are Siamese Twins", CEPR Discussion Paper No. 3309.

Bovenberg, A. Lans, and Frederick van der Ploeg (1994). "Effects of the Tax and Benefit System on Wage Formation and Unemployment", mimeo: Tilburg University.

Cameron, Stephen, and James J. Heckman (1999). "Can tuition Policy Combat Inequality?", in: M. Kosters, ed., *Financing Tuition: Government Policies and Social Priorities*, Washington D.C.: AEI Press.

Cameron, Stephen, and Christopher Taber (2000). "Borrowing Constraints and the Returns to Schooling", NBER Workingpaper 7761.

Corcoran, Mary, Roger Gordon, Deborah Laren, and Gary Solon (1992). "The Association Between Men's Economic Status and Their Family Origins", *Journal of Human Resources*, 27, (4), 575-601.

Diamond, Peter A., and James A. Mirrlees (1971). "Optimal Taxation and Public Production I: Production Efficiency", *American Economic Review*, 61, 8-27.

Dixit, Avinash K., and Agnar Sandmo (1977). "Some Simplified Formulae for Optimal Income Taxation", *Scandinavian Journal of Economics*, 79, 417-423.

Durlauf, Steven N. (1996). "A Theory of Persistent Income Inequality", *Journal of Economic Growth*, 1, 75-93.

Eaton, Jonathan and Harvey S. Rosen (1980). "Taxation, Human Capital and Uncertainty", *American Economic Review*, 70, (4), 705-715.

Ewijk, Casper van, and Paul J.G. Tang (2000). "Efficient Progressive Taxes and Education Subsidies", CPB Research Memorandum 170, CPB Netherlands Bureau for Economic Policy Analysis.

Fernàndez, Raquel, and Richard Rogerson (1996). "Income Distribution, Communities, and the Quality of Public Education", *Quarterly Journal of Economics*, 111(1), 135-164.

Fernàndez, Raquel, and Richard Rogerson (1998). "Public Education and Income Distribution: A Quantitative Evaluation of Education Finance Reform", *American Economic Review*, 88, 813-833.

Friedman, Milton (1962). *Capitalism and Freedom*, Chicago: Chicago University Press.

Friedman, Milton, and Simon Kuznets (1945). *Income from Independent Professional Practice*, New York: National Bureau for Economic Research.

Galor, Oded, and Joseph Zeira (1993). "Income Distribution and Macroeconomics", *Review of Economic Studies*, 60, 35-52.

Haveman, Robert, and Barbara Wolfe (1995). "The Determinants of Children's Attainments: A Review of Methods and Findings", *Journal of Economic Literature*, 33, 1829-1878.

Heckman, James J. (1976). "A Life-Cycle Model of Earnings, Learning and Consumption", *Journal of Political Economy*, 4, S11-S44.

Heckman, James J., Lance Lochner, and Christopher Taber (1998). "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents", *Review of Economic Dynamics*, 1, 1-58.

Hubbard, R. Glenn, and Kenneth L. Judd (1986). "Liquidity Constraints, Fiscal Policy, and Consumption", *Brookings Papers on Economic Activity*, 1, 1-59.

Jacobs, Bas (2001). "Optimal Income Taxation with Endogenous Human Capital", mimeo: University of Amsterdam.

Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi (1993). "Optimal Taxation in Models of Endogenous Technological Change", *Journal of Political Economy*, 101 (3), 485-517.

Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi (1997). "On the Optimal Taxation of Capital Income", *Journal of Economic Theory*, 73, 93-117.

Judd, Kenneth L. (1999). "Optimal Taxation and Spending in General Competitive Growth Models", *Journal of Public Economics*, 71, 1-26.

Kane, Thomas J. (1994). "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education", *Journal of Political Economy*, 102 (5), 878-911.

Kremer, Michael (1997). "How Much Does Sorting Increase Inequality?", *Quarterly Journal of Economics*, 113, 115-139.

Lazear, Edward P. (1983). "Intergenerational Externalities", *Canadian Journal of Economics*, 16 (2), 212-228.

Loury, Glenn C. (1981). "Intergenerational Transfers and the Distribution of Earnings", *Econometrica*, 49, (4), 843-867.

Milesi-Ferretti, Gian Maria, and Nouriel Roubini (1998). "On the Taxation of Human and Physical Capital in Models of Endogenous Growth", *Journal of Public Economics*, 70, 237-254.

Mincer, Jacob (1962). "On the Job Training: Costs, Returns and Some Implications", *Journal of Political Economy*, 70, 50-79.

Mirrlees, James A. (1971). "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, 38, 175-208.

Nielsen, Søren B., and Peter B. Sørensen (1997). "On the Optimality of the Nordic System of Dual Income Taxation", *Journal of Public Economics*, 63, 311-329.

Perotti, Roberto (1993). "Political Equilibrium, Income Distribution, and Growth", *Review of Economic Studies*, 60, 755-776.

Plug, Erik J. S., and Wim Vijverberg (2001). "Schooling, Family Background, and Adoption: Does Family Income Matter?", mimeo: University of Amsterdam.

Shea, John (2000). "Does Parents' Money Matter?", *Journal of Public Economics*, 77, 155-184.

Sørensen, Peter B. (1999). "Optimal Tax Progressivity in Imperfect Labor Markets", *Labor Economics*, 6, 435-452.

Stern, Nicholas H. (1976). "On the Specification of Models of Optimum Income Taxation", *Journal of Public Economics*, 6, 123-162.

Stiglitz, Joseph E. and Andrew Weiss (1981). "Credit Rationing in Markets with Imperfect Information", *American Economic Review*, 71, (3), 393-410.

Stokey, Nancy L., and Sergio T. Rebelo (1995). "Growth Effects of Flat-Rate Taxes", *Journal of Political Economy*, 103, 519-550.

Trostel, Philip A. (1993). "The Effect of Taxation on Human Capital", *Journal of Political Economy*, 101, 327-350.

Tuomala, Matti (1985). "Simplified Formulae for Optimal Linear Income Taxation", *Scandinavian Journal of Economics*, 87 (4), 668-672.

Tuomala, Matti (1990). *Optimal Income Tax and Redistribution*, Oxford: Clarendon Press.

Weiss, Yoram. (1986). "The Theory of Life-Cycle Earnings", in: Ashenfelter, Orley and Richard Layard (1986). *Handbook of Labor Economics.* Amsterdam: Elsevier Science Publishers BV.

Appendix

Household optimization

Households maximize their utility function subject to the budget constraints in both periods and the production function for human capital. The Lagrangian for the problem is given by - suppressing the agent's indices α and ω:

$$\begin{aligned}\mathcal{L} = u(c_1, c_2) &+ \lambda_1((1-t)w(1-x) + \omega + g - \kappa x - c_1 - s) \\ &+ \lambda_2((1-t)w\phi(.) + (1+r)s + g - c_2) - \mu s,\end{aligned}$$

where λ_1, λ_2, and μ denote the Lagrange multipliers of the first and the second period budget constraint, and the non-negativity constraint on savings respectively. First-order conditions (FOC's) are given by:

$$\begin{aligned}
&\frac{\partial \mathcal{L}}{\partial c_1} = u_{c_1} - \lambda_1 = 0, \\
&\frac{\partial \mathcal{L}}{\partial c_2} = u_{c_2} - \lambda_2 = 0, \\
&\frac{\partial \mathcal{L}}{\partial x} = -\lambda_1((1-t)w + \kappa) + \lambda_2(1-t)w\phi'(.) = 0, \\
&\frac{\partial \mathcal{L}}{\partial s} = -\lambda_1 - \mu + (1+r)\lambda_2 \leq 0, \quad s \geq 0, \quad s\frac{\partial \mathcal{L}}{\partial s} = 0, \\
&\frac{\partial \mathcal{L}}{\partial \lambda_1} = \frac{\partial \mathcal{L}}{\partial \lambda_2} = 0, \\
&\frac{\partial \mathcal{L}}{\partial \mu} \geq 0, \quad \mu \geq 0, \quad \mu\frac{\partial \mathcal{L}}{\partial \mu} = 0,
\end{aligned}$$

where we assumed implicitly that $c_1, c_2 \geq 0$. The first two FOC's state that marginal utility of consumption in both periods should equal their marginal

value as reflected in shadow prices λ_1 and λ_2. The third FOC that marginal costs and benefits of learning must be equalized in an optimal solution. In the fourth FOC and last FOC, the Kuhn-Tucker conditions are given for an optimal choice of s, where either savings are 0 if the constraint is binding - the marginal value of a relaxation of the constraint, as measured by μ, is positive - or savings are positive if the constraint is non-binding - the marginal value of a relaxation of the constraint is zero, i.e. $\mu = 0$. The complementary slackness condition, the last term, captures both these elements for an optimal choice. The fifth FOC restates that the two budget constraints must hold as strict equalities for an optimal choice.

We can rearrange the FOC to arrive at the marginal rate of inter-temporal substitution in consumption:

$$\frac{u_{c_1}}{u_{c_2}} = \frac{\lambda_1}{\lambda_2}.$$

Inversion of this equation yields the Frisch demand equations given in the text. The demand for inputs invested in education follows from inversion of:

$$\frac{(1-t)w\phi'(.)}{(1-t)w+\kappa} = \frac{\lambda_1}{\lambda_2} \geq 1+r.$$

Derivation of partial derivatives $\frac{\partial H}{\partial G}$ and $\frac{\partial H}{\partial t}$

First, note that $\frac{\partial H}{\partial t} = \frac{\partial H}{\partial x}\frac{\partial x}{\partial t}$ and $\frac{\partial H}{\partial G} = \frac{\partial H}{\partial x}\frac{\partial x}{\partial G}$. From the definition for $H \equiv w(1-x) + w\phi(\alpha; x)/(1+r)$ follows:

$$\frac{\partial H}{\partial x} = w\left(\frac{\phi_x(\alpha;x)}{1+r} - 1\right) = w\left(\frac{\lambda_1/\lambda_2}{1+r}\left(1+\frac{\kappa}{w(1-t)}\right) - 1\right) > 0,$$

where we substituted the first order condition for learning. Since $\lambda_1/\lambda_2 \geq 1+r$ and $\kappa > 0$ the first term in brackets is always larger than 1. Next we need to determine the signs of $\frac{\partial x}{\partial t}$ and $\frac{\partial x}{\partial G}$.

Unconstrained consumers

The first-order condition for learning is:

$$(1-t)w\phi_x(\alpha, x) = (1+r)((1-t)w+\kappa).$$

It is immediately clear that $\frac{\partial x}{\partial G} = 0$. Linearizing the first-order condition yields:

$$\tilde{x} = -\frac{\omega_\kappa}{\varepsilon_{\phi\phi}}\tilde{t},$$

where $\tilde{x} \equiv dx/x$ and $\tilde{t} \equiv dt/(1-t)$ denote relative changes in learning time and the tax rate, $\varepsilon_{\phi\phi} \equiv -x\phi''/\phi' > 0$ is the elasticity of marginal productivity of learning, and $\omega_\kappa \equiv \left(\frac{\kappa}{w(1-t)+\kappa}\right) > 0$ is the share of direct costs in total costs of education. From the fact that $\tilde{x}/\tilde{t} < 0$, follows that $\partial x/\partial t < 0$, and, hence, $\frac{\partial H}{\partial t} < 0$.

Constrained consumers

Here, the derivation is more complicated. The first-order condition for learning is given by:

$$(1-t)w\phi_x(\alpha, x) = \frac{u_{c_1}}{u_{c_2}}(w(1-t)+\kappa).$$

Combined with the two budget constraints the level of x is determined. Linearizing the first order condition yields:

$$\frac{1}{\sigma}(\tilde{c}_2 - \tilde{c}_1) = -\varepsilon_{\phi\phi}\tilde{x} - \omega_\kappa\tilde{t},$$

where $\sigma \equiv \frac{d\ln(c_2/c_1)}{d\ln(u_{c_1}/u_{c_2})}$ is the inter-temporal elasticity of substitution in consumption, and $\tilde{c}_1 \equiv dc_1/c_1$, and $\tilde{c}_2 \equiv dc_2/c_2$ denote the relative changes in consumption. Linearizing the first period budget constraint gives:

$$\tilde{c}_1 = -\gamma_1\tilde{t} - \gamma_x\tilde{x} + \gamma_{g1}\tilde{g},$$

where $\gamma_1 \equiv (1-t)w(1-x)/c_1$ is the share of first period labor income in total income, $\gamma_x \equiv ((1-t)w+\kappa)x/c_1$ is the share of outlays on education in first period income, $\gamma_{g1} \equiv g/c_1$ is the share of the lump-sum transfer in first period income, and $\tilde{g} \equiv dg/g$. Note that $\gamma_1 + \gamma_x + \gamma_{g1} + \gamma_\omega = 1$, where $\gamma_\omega \equiv \omega/c_1$ is the share of the transfer in first period consumption. The linearized second period budget constraint is given by:

$$\tilde{c}_2 = \gamma_2(\varepsilon_\phi\tilde{x} - \tilde{t}) + \gamma_{g2}\tilde{g},$$

where $\varepsilon_\phi \equiv x\phi'/\phi$ is the elasticity of the production function for human capital, $\gamma_2 \equiv (1-t)w\phi(.)/c_2$ is the share of second period income in second

period consumption and $\gamma_{g2} \equiv g/c_2$ denotes the share of the transfer in second period consumption ($\gamma_2 + \gamma_{g2} = 1$). Solving for $\tilde{x}$ gives:

$$\tilde{x} = -\left(\frac{\gamma_1 - \gamma_2 + \sigma\omega_\kappa}{\gamma_2\varepsilon_\phi + \sigma\varepsilon_{\phi\phi} + \gamma_x}\right)\tilde{t} + \left(\frac{\gamma_{g1} - \gamma_{g2}}{\gamma_2\varepsilon_\phi + \sigma\varepsilon_{\phi\phi} + \gamma_x}\right)\tilde{G}.$$

where $\tilde{G} = \tilde{g}$. It is clear that $\frac{\partial x}{\partial G} > 0$ in this case, since $\left.\frac{\tilde{x}}{\tilde{G}}\right|_{\tilde{t}=0} > 0$ due to $\gamma_{g1} - \gamma_{g2} > 0$ because second period consumption is larger than first period consumption. Therefore, $\frac{\partial H}{\partial G} > 0$. The sign of $\frac{\partial x}{\partial t}$ is derived as follows. First note that $\gamma_1 - \gamma_2 = \gamma_{g_2} - (1 - \gamma_1)$. We assume that $\gamma_1 - \gamma_2 + \sigma\omega_\kappa > 0$. There is a positive effect of taxation on learning since the share of second period income is higher than first period income, $\gamma_1 < \gamma_2$. The intuition for this is effect is that the marginal value of second period income is more affected by taxation, than the marginal value of first-period income. λ_1/λ_2 decreases, and more time is spend on learning. We assume that the first effect dominates the second effect, otherwise the problem is ill defined (both taxation and the transfer exert a positive effect on learning). Consequently, $\frac{\partial H}{\partial t} < 0$.

Derivation optimal tax formula with heterogeneous agents

The first order conditions for the government are:

$$\frac{\partial \mathcal{L}}{\partial G} = \int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\left(\Psi'\frac{\partial V}{\partial G} - \eta + \eta t\frac{\partial H}{\partial G}\right)dF = 0,$$

$$\frac{\partial \mathcal{L}}{\partial t} = \int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\left(\Psi'\frac{\partial V}{\partial t} + \eta H + \eta t\frac{\partial H}{\partial t}\right)dF = 0.$$

Use Roy's lemma to establish that $\frac{\partial V}{\partial G} = \lambda_1$ and $\frac{\partial V}{\partial t} = -\lambda_1(w(1-x)) - \lambda_2 w\phi(\alpha; x)$. The last expression can be rewritten so as to get:

$$\frac{\partial V}{\partial t} = \lambda_1\frac{w\phi(\alpha; x)}{1+r} - \lambda_1 H + \lambda_2 w\phi(\alpha; x) = -\lambda_1 H(1-q),$$

where $\chi \equiv w\phi(\alpha; x)/H$ is the share of second period income total income, and $q \equiv \chi\left(\frac{1}{1+r} - \frac{\lambda_2}{\lambda_1}\right) > 0$ measures the welfare loss due to the capital

market imperfection. Substitution in the first order condition for G yields the marginal costs of public funds ($MCPF \equiv \eta / \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 dF$):

$$MCPF = \frac{\eta}{\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 dF} = \frac{1}{1 - \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} t \frac{\partial H}{\partial G} dF} > 1.$$

Substitution of Roy's lemma in the first-order condition for the optimal tax rate gives:

$$\begin{aligned} & -\frac{t}{1-t} \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \frac{(1-t)}{H} \frac{\partial H}{\partial t} H dF \\ & = -\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \frac{\Psi' \lambda_1 H}{\eta} dF + \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} H dF + \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \frac{\Psi' \lambda_1 H q}{\eta} dF. \end{aligned}$$

Define the average weighted elasticity of life-time earnings w.r.t. the tax rate as follows:

$$\bar{\varepsilon} \equiv \frac{-\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \frac{(1-t)}{H} \frac{\partial H}{\partial t} H dF}{\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} H dF} > 0.$$

Further, define the normalized covariance between the marginal social value of income $\lambda = \Psi' \lambda_1$and earnings H:

$$\begin{aligned} \xi &\equiv \frac{\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 dF \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} H dF - \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 H dF}{\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 dF \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} H dF} \\ &= \frac{-cov[\Psi' \lambda_1, H]}{\int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} \Psi' \lambda_1 dF \int_{\underline{\alpha}}^{\infty} \int_{\underline{\omega}}^{\infty} H dF} > 0. \end{aligned}$$

Due to the concavity of utility and social welfare functions, the covariance is negative so that the distributional characteristic $\xi > 0$, see Dixit and Sandmo (1977).

Using these definitions, we can rewrite the expression for t:

$$
\begin{aligned}
&-\frac{t}{1-t}\bar{\varepsilon} = 1 - \frac{\int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\Psi'\lambda_1}{\eta}HdF}{\bar{H}} + \frac{\int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\Psi'\lambda_1}{\eta}HqdF}{\bar{H}} \\
&= \frac{\lambda\bar{H}}{\lambda\bar{H}} - \frac{\int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\Psi'\lambda_1 HdF + (\eta-\lambda)\int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\Psi'\lambda_1}{\eta}HdF}{\lambda\bar{H}} \\
&+ \frac{\int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\Psi'\lambda_1}{\eta}HqdF}{\bar{H}} \\
&= \frac{\lambda\bar{H} - \int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\Psi'\lambda_1 HdF}{\lambda\bar{H}} + \frac{(MCPF-1)\int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\Psi'\lambda_1}{\eta}HdF}{\bar{H}} \\
&+ \frac{\int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\Psi'\lambda_1 Hq}{\eta}dF}{\bar{H}} \\
&= \frac{-cov[\Psi'\lambda_1, H]}{\lambda\bar{H}} + \frac{(MCPF-1)}{\overline{MCPF}} + \overline{\left(\frac{q}{MCPF}\right)},
\end{aligned}
$$

where $\bar{H} \equiv \int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty} H_{\alpha\omega}dF(\alpha,\omega)$, $\frac{1}{\overline{MCPF}} \equiv \int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\Psi'\lambda_1}{\eta}HdF/\bar{H} > 0$ is the income weighted inverse of $MCPF$, and $\overline{\left(\frac{q}{MCPF}\right)} \equiv \int_{\underline{\alpha}}^{\infty}\int_{\underline{\omega}}^{\infty}\frac{\Psi'\lambda_1}{\eta}qHdF/\bar{H} > 0$ is the income weighted marginal social welfare loss associated with the capital market imperfection. The optimal tax rate satisfies:

$$
\frac{t}{1-t} = \frac{1}{\bar{\varepsilon}}\left(\xi + \frac{(MCPF-1)}{\overline{MCPF}} + \overline{\left(\frac{q}{MCPF}\right)}\right).
$$

Solution Cobb-Douglas case

The solution of the pair of endogenous variables (c_1, c_2, x, s) for the unrestricted agents ($s \geq 0$) is (note that $g = G$ since g is set at 0 in the second period):

$$
\begin{aligned}
x^u &= \left(\frac{\gamma A\alpha^{\beta}}{(1+r)(1+\kappa/(1-t)w)}\right)^{\frac{1}{1-\gamma}}, \\
s^u &= \frac{1}{2}((1-t)w+\omega+G) - \frac{1}{2\gamma}(1+\gamma)((1-t)w+\kappa)x^u, \\
c_1^u &= \frac{1}{2}((1-t)w+\omega+G) + \frac{1}{2\gamma}(1-\gamma)((1-t)w+\kappa)x^u, \\
c_2^u &= \frac{1}{2}(1+r)((1-t)w+\omega+G) + \frac{1}{2\gamma}(1-\gamma)(1+r)((1-t)w+\kappa)x^u,
\end{aligned}
$$

where a superscript u denotes a solution in the *un*constrained case. Further, the solution (c_1, c_2, x, s) for the constrained agents is:

$$
\begin{aligned}
x^c &= \frac{\gamma}{1+\gamma}\left(\frac{(1-t)w+\omega+G}{(1-t)w+\kappa}\right), \\
s^c &= 0, \\
c_1^c &= \frac{(1-t)w+\omega+G}{1+\gamma}, \\
c_2^c &= (1-t)wA\alpha^\beta(x^c)^\gamma,
\end{aligned}
$$

where a superscript c denotes a solution in the *con*strained case.