WIF - Institute of Economic Research

Tax Policy and Human Capital Formation with Public Investment in Education

Simone Valente

Working Paper 05/41

June 2005

Economics Working Paper Series

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Tax policy and human capital formation with public investment in education

Simone Valente*

June 2005

Abstract

This paper studies the effects of distortionary taxes and public investment in an endogenous growth OLG model with knowledge transmission. Fiscal policy affects growth in two respects: first, work time reacts to variations of prospective tax rates and modifies knowledge formation; second, public spending enhances labour efficiency but also stimulates physical capital through increased savings. It is shown that Ramsey-optimal policies reduce savings due to high tax rates on young generations, and are not necessarily growth-improving with respect to a pure private system. Non-Ramsey policies that shift the burden on adults are always growth-improving due to crowding-in effects: the welfare of all generations is unambiguously higher with respect to a private system, and there generally exists a continuum of non-optimal tax rates under which long-run growth and welfare are higher than with the Ramsey-optimal policy.

JEL classification: E62, O41, O11.

Keywords: Endogenous growth, Human capital, Overlapping generations, Tax policy, Public investment.

*Institute of Economic Research, ETH Zurich. Address: Zurichbergstrasse 18 - ZUE F15, CH-8032 Zurich (Switzerland). E-mail: svalente@ethz.ch.

1 Introduction[1]

After Lucas' (1988) seminal contribution, the view that human capital formation drives economic development inspired a huge body of literature on endogenous growth. Several authors investigated the sources of knowledge accumulation at both the theoretical and empirical levels, emphasising the role of educational attainment and knowledge spillovers in raising aggregate productivity (Benhabib and Spiegel, 1994). Empirical evidence suggests that monetary investments in education, and public spending in particular, are also relevant in determining the accumulation rate of human capital (Barro and Sala-i-Martin, 1995).[2] However, the link between taxation, public investment in education, and knowledge transmission, has not been fully analysed at the theoretical level. Several recent studies analyse the growth effects of taxation in Lucas-type settings with overlapping generations, where knowledge transmission determines an intergenerational externality. This literature typically assumes that *study time* affects knowledge through a learning process: given a finite amount of time to be allocated between studying and working, individual choices determine a tradeoff between human and physical capital accumulation at the aggregate level (see *e.g.* De Gregorio, 1996; Yakita, 2003). In fact, some recent contributions also include *educational expenditures* in the learning process, and analyse the long-run effects of alternative policies by means of simulations using computable general equilibrium models (Docquier and Michel, 1999; Hendricks, 1999; Bouzahzah *et al.,* 2002). However, most analytical results on the growth effects of taxation rule out productive educational expenditures: Bovenberg and van Ewijk (1997), Meijdam (1998), Heijdra and Ligthart (2000), and Yakita (2003) examine long-run distortions induced by taxing capital and labour incomes without assuming labour-enhancing monetary investment.[3]

This paper studies the effects of alternative tax policies when both study

[1] I thank CentER, Tilburg University, and CeFiMS, University of London, for hospitality, and Barbara Annicchiarico, Fabrizio Adriani, Christa Brunnschweiler, Luca Deidda, Giancarlo Marini, Pasquale Scaramozzino, Sjak Smulders, and three anonymous Referees for comments and suggestions. Financial support from the European Commision - Marie Curie Fellowship - is gratefully acknowledged.

[2] According to the cross-country analysis by Barro and Sala-i-Martin (1995, Chap.13), a 1.5 percent increase in the public education spending-GDP ratio would have raised the average growth rate by 0.3 percent per year in the period 1965-1975.

[3] An exception is Buiter and Kletzer (1995): the authors study the effects of borrowing constraints when private and public spending in education are perfect substitutes.

time and educational expenditure improve human capital formation. In this setting, fiscal policy influences long-run growth through two channels: (i) the reallocation of time between studying and working induced by variations of prospective tax rates, and (ii) the link between educational spending, private savings and knowledge accumulation. With respect to Bovenberg and van Ewijk (1997), Meijdam (1998), and Yakita (2003), mechanism (ii) is peculiar to this model, since public spending may imply substantial *crowding-in effects*. It is shown that these two mechanisms imply an inverted-U relation between long-run growth and the second-period tax rate. The reason for this result is that the amount of study time and the asymptotic savings rate react in opposite ways to balanced variations of tax rates: taxing young individuals increases study time but lowers savings by reducing disposable income; conversely, taxing adults boosts physical capital accumulation. The implication of the inverted-U relation is that any growth rate (except the unique maximum level) is associated with opposite financing strategies: 'taxing the young' or 'taxing adults'.

Building on this result, we analyse alternative tax policies and discuss their implications in terms of growth, welfare and intergenerational equity. The typical starting point is to characterise Ramsey-optimal policies, *i.e.* policies that implement the intertemporal allocation which maximises social welfare, according to the standard utilitarian social welfare function. It is shown that Ramsey-optimal policies imply high taxation in the first period of life, and may be growth-improving with respect to a pure private system depending on parameter values. A crucial role is played by the relative share of physical capital in production: high tax rates on young generations reduce savings, implying higher (lower) growth rates and welfare levels in the long run if crowding-out effects are weak (strong) enough.

Considering alternative tax rules, it is shown that crowding-in effects are important when the government pursues the opposite financing strategy, i.e. 'taxing adults': shifting the burden from the first to the second period of life allows sustaining growth and welfare through increased savings. In particular, tax rates may be adjusted so that work time of young agents is the same as without public intervention: under this *labour-neutral* policy, growth and welfare are always higher with respect to a pure private system, irrespective of the parameter values. As a consequence, labour-neutral policies may increase long-run growth and preserve welfare of later generations when compared with Ramsey-optimal policies.

Furthermore, the inverted-U relation between growth and second-period tax rates allows us to characterise a *growth-equivalent policy* under which the asymptotic growth rate equals that obtained under Ramsey-optimal taxation. This implies that there generally exists a set of policies yielding higher growth and increased saving rates in the long run than those observed under Ramsey-optimal taxation.

2 The model

The analysis employs an overlapping-generations model where human capital accumulation is enhanced by monetary investment in education. In order to assess the effects of fiscal policy more neatly, the dynamic path experienced in a pure *public regime* - where education is entirely financed by the government - is compared with that obtained in a pure *private regime*, where young individuals pay their own education costs. Although the analysis is connected with the literature on growth and education financing (Glomm and Ravikumar, 1992; Eckstein and Zilcha, 1994; De Gregorio and Kim, 2000), the aim of this paper is not to discuss the desirability of a particular school system, but rather to describe the intergenerational consequences of alternative financing strategies within the public regime: the private system will be thought of as a comparable benchmark economy, by means of which the implications of different tax rules for growth and welfare can be discussed.

Consider two economies indexed by $i = A$, B, with identical technologies, preferences, initial endowments, and a constant population of consumer-workers who live for two periods. In period t there are n young and n adult individuals, and each young individual inherits own knowledge from the current state of the economy. Knowledge is represented by $\bar{h}^i$, measured in terms of labour-efficiency units. Individuals are endowed with one unit of time: in the first period of life a fraction $(1-\ell_t^i)$ is devoted to study, and $\ell_t^i \bar{h}_t^i$ labour units are supplied for production. In the second period, individuals only work, and consume all their income.[4] The level of efficiency achieved at the beginning of the second period of life depends on study time and school

[4]A relevant issue not addressed in this paper is the interplay between education financing and pension funding, which would require including a third period of life in which agents only consume. While a three-period version would require numerical solutions, the basic two-period setup employed here allows to obtain analytical results on crowding-in effects, filling the gap in previous literature. Extending the model to include pensions is nonetheless an interesting topic which deserves further research.

quality E^i, according to the learning technology

$$\bar{h}^i_{t+1} = \bar{h}^i_t \varphi^i_t, \tag{1}$$

$$\varphi^i_t = \Psi \left(1-\ell^i_t\right)^{\varepsilon} \left(E^i_t\right)^{\eta}, \qquad i = A, B, \tag{2}$$

where $\Psi > 0$ is a proportionality factor, φ exhibits decreasing returns in both arguments ($0 < \varepsilon < 1$, $0 < \eta < 1$) and non-increasing returns to scale ($\varepsilon + \eta \leq 1$). Aggregate human capital H is the amount of labour supplied by the two generations alive in period t: denoting by $h^i_t = n\bar{h}^i_t$ the aggregate amount of knowledge in each generation, human capital at time t equals $H^i_t = \left(1+\ell^i_t\right) h^i_t$. Since agents have identical preferences, total labour supply evolves according to

$$H^i_{t+1} = \left(1+\ell^i_{t+1}\right) h^i_t \varphi^i_t = \varphi^i_t H^i_t \left(1+\ell^i_{t+1}\right) \left(1+\ell^i_t\right)^{-1}. \tag{3}$$

Aggregate output (Y) is produced by means of human and physical capital (K) according to the production function $Y = K^{\alpha} H^{1-\alpha}$, with $0 < \alpha < 1$. Physical capital fully depreciates during the production process. This widely-used assumption is relevant for the analysis: on the one hand, assuming one-period depreciation of K is reasonable in the present model - since a 'period' corresponds to one half of the individual life cycle - and it allows us to study dynamics analytically; on the other hand, the asymmetric treatment of physical and human capital does not necessarily weaken the main results - namely, Propositions 4 and 6 below - since assuming durable physical capital might emphasise the role of crowding-in effects in the long run.

Setting $k = K/H$, the output-human capital ratio $y = Y/H$ equals

$$y^i_t = \left(k^i_t\right)^{\alpha}. \tag{4}$$

The production sector behaves like a single competitive firm: denoting by w and R the wage rate and the interest factor respectively, profit maximisation implies

$$R^i_t = \alpha \left(Y^i_t / K^i_t\right), \tag{5}$$

$$w^i_t = (1-\alpha) \left(k^i_t\right)^{\alpha}, \quad i = A, B, \tag{6}$$

Individual consumption is denoted by c when young, and by d when adult. Preferences are logarithmic and individual lifetime utility U is

$$U^i_t = \log \left(c^i_t\right) + \beta \log \left(d^i_{t+1}\right), \tag{7}$$

where $\beta \in (0,1)$ is the private discount factor.

School quality is indexed by the levels of private and public spending in education, which are assumed to be equally productive. Economy A is a pure *private system* where education costs V are paid by young generations and there is no public intervention: setting $v = V/h$, school quality equals $E_t^A = v_t$. Economy B is a pure *public* school system with total spending in education G financed through distortionary taxation: setting $g = G/h$ yields $E_t^B = g_t$. In the present context, assuming that private and public spending are equally productive is formally equivalent to considering two extreme cases of the general learning technology $\varphi = \Psi (1-\ell)^{\varepsilon} (v+g)^{\eta}$, which represents a mixed school system with v and g as perfect substitutes.[5]

Firstly, consider the temporary equilibrium in the private system. In economy A, each consumer faces the budget constraints

$$\begin{aligned} c_t^A &= w_t^A \ell_t^A \bar{h}_t^A - v_t \bar{h}_t^A - s_t^A, & (8)\\ d_{t+1}^A &= R_{t+1}^A s_t^A + w_{t+1}^A \bar{h}_{t+1}^A, & (9) \end{aligned}$$

where s represents individual savings. Each agent maximises U_t^A subject to (8)-(9), using $c_t^A, d_{t+1}^A, \ell_t^A, v_t$ as control variables, taking w and R as given. First-order conditions read

$$\begin{aligned} R_{t+1}^A w_t^A &= -\varphi_{\ell_t}^A w_{t+1}^A, & (10)\\ R_{t+1}^A &= \varphi_{v_t}^A w_{t+1}^A, & (11)\\ d_{t+1}^A &= \beta c_t^A R_{t+1}^A. & (12) \end{aligned}$$

Denote aggregate savings by $S_t = ns_t$ and set $S_t^A \equiv K_{t+1}^A$. Substituting equilibrium prices (5)-(6) and the first-order conditions in individual budget constraints gives the accumulation rule

$$k_{t+1}^A = \frac{(1-\alpha)\left[\ell_t^A (1+\beta\eta+\beta\varepsilon) - 1 - \beta\eta\right]}{\varepsilon (1+\beta)\left(1+\ell_{t+1}^A\right)\varphi_t^A} y_t^A. \tag{13}$$

The optimal amount of work time supplied by young generations determines, together with (13), the temporary equilibrium of the economy, which is de-

[5] Assuming that v and g are equally productive is not particularly restrictive here. Since we analyse alternative public policies using the private system as a comparable benchmark economy, ruling out perfect substitutability would "[...] create a role for government in human capital formation that is too straightforward" because it "would add to the algebra without qualitatively changing the effects of public spending on human capital formation and *private financial saving*" (Buiter and Kletzer, 1995, p.S168; our italics).

fined at given expectations over the future interest rate and the future employment level. When there is a tradeoff between studying and working, it is possible to obtain *stationary solutions*, where work time jumps at the equilibrium level $\ell_\star$ in period zero and is constant thereafter (De Gregorio, 1996; de la Croix and Michel, 2002). In our model, the assumed learning technology (2) implies a stationary solution (all proofs are in the Appendix):

Lemma 1 *In the private education regime, work time supplied by young generations is equal to the optimal level $\ell_\star^A$ in each period, with*

$$\begin{aligned} \ell_\star^A &= (1/2)\left[1-\varepsilon q^A - p^A + \sqrt{\left(1-\varepsilon q^A - p^A\right)^2 + 4p^A}\right], &(14)\\ \ell_\star^A &> \ell_{\min}^A = \frac{1+\beta\eta}{1+\beta\varepsilon+\beta\eta}. &(15)\end{aligned}$$

Coefficients $q^A > 0$ and $p^A > 1$ in eq.(14) are constant parameters, and the lower bound $\ell_{\min}^A$ is the minimum amount of work time compatible with positive savings - see eq.(13). Denoting the *private propensity to spend in education* as $\rho_t^A = (v_t/y_t^A)$, condition (11) can be rewritten as

$$v_t = \rho_t^A y_t^A = \frac{\eta}{\varepsilon}(1-\alpha)\left(1-\ell_\star^A\right) y_t^A. \tag{16}$$

Hence, the optimal propensity is time-invariant, and the accumulation rule (13) may be rewritten as

$$k_{t+1}^A = z^A \left(y_t^A\right)^{1-\eta} = z^A \left(k_t^A\right)^{\alpha(1-\eta)}, \tag{17}$$

where the *accumulation rate* z is constant and equal to

$$z^A = \frac{(1-\alpha)\left[\ell_\star^A (1+\beta\eta+\beta\varepsilon) - 1 - \beta\eta\right]}{\varepsilon(1+\beta)\left(1+\ell_\star^A\right)\Psi\left(1-\ell_\star^A\right)^\varepsilon \left(\rho^A\right)^\eta}. \tag{18}$$

Since $\alpha(1-\eta) < 1$, the physical-human capital ratio converges to a steady-state level in the long run.

In the public regime (economy B), the government imposes proportional taxes on labour incomes, and individual budget constraints read

$$\begin{aligned} c_t^B &= w_t^B \ell_t^B \bar{h}_t^B (1-x_t) - s_t^B, &(19)\\ d_{t+1}^B &= R_{t+1}^B s_t^B + w_{t+1}^B \bar{h}_{t+1}^B (1-\theta_{t+1}), &(20)\end{aligned}$$

where x and θ are proportional tax rates on labour earnings in the first and in the second period of life, respectively.[6] The government keeps a balanced budget in each period:

$$g_t h_t^B = w_t^B h_t^B \left(\theta_t + x_t \ell_t^B\right). \tag{21}$$

Individuals anticipate tax rates with perfect foresight and maximise U_t^B subject to (19)-(20) using c_t^B, d_t^B, and ℓ_t^B as control variables. First-order conditions are

$$\begin{aligned} R_{t+1}^B w_t^B (1 - x_t) &= -\varphi_{\ell_t}^B w_{t+1}^B (1 - \theta_{t+1}), &(22)\\ d_{t+1}^B &= \beta c_t^B R_{t+1}^B. &(23) \end{aligned}$$

Since there is no public debt, net investment equals aggregate savings and the accumulation rule of the economy is

$$k_{t+1}^B = \frac{(1-\alpha)\left[\ell_t^B (1+\beta\varepsilon) - 1\right](1 - x_t)}{\varepsilon (1+\beta)\left(1 + \ell_{t+1}^B\right)\varphi_t^B} y_t^B. \tag{24}$$

Substituting condition (22) in (24) yields

$$\begin{aligned} \ell_{t+1}^B &= \frac{\ell_t^B}{1 - \ell_t^B} \varepsilon q_{t+1}^B - p_{t+1}^B, &(25)\\ q_{t+1}^B &= \beta (1-\alpha)(\alpha + \alpha\beta)^{-1} (1 - \theta_{t+1}), &(26)\\ p_{t+1}^B &= 1 + (1-\alpha)(\alpha + \alpha\beta)^{-1} (1 - \theta_{t+1}). &(27) \end{aligned}$$

Expressions (25)-(26)-(27) show that ℓ^B depends on prospective tax rates. In particular, when θ_t is kept constant, public regimes also exhibit stationary work time:

[6] On the one hand, allowing tax rates to differ between generations ensures that the Ramsey-optimal allocation can be implemented through instruments x, ρ and θ as shown in section 3.1. On the other hand, x and θ may be interpreted in terms of net marginal burdens without loss of generality: this reflects the possibility for the government to implement age-uniform flat tax rates on income together with public policies that modify *ex-post* fiscal wedges. These intergenerational distortions arise, for example, when the government combines income taxes with subsidies *à la* Docquier and Michel (1999), *i.e.* monetary subsidies that relieve the opportunity cost of studying when young agents seek higher education and/or skill-enhancing training.

Lemma 2 *If $\theta_t = \theta$ in each period, work time of young generations in the public regime equals $\ell_\star^B$ in each period, where*

$$\ell_\star^B = (1/2)\left[1 - \varepsilon q^B - p^B + \sqrt{\left(1 - \varepsilon q^B - p^B\right)^2 + 4p^B}\right], \tag{28}$$

$$\ell_\star^B > \ell_{\min}^B = \frac{1}{1+\beta\varepsilon}. \tag{29}$$

When both tax rates are time-invariant, the economy converges to the balanced growth path: denoting the *public propensity to spend in education* by $\rho_t^B = (g_t/y_t^B)$ and setting $x_t = x$ and $\theta_t = \theta$ in the government budget constraint, the accumulation rule (24) becomes

$$k_{t+1}^B = z^B \left(y_t^B\right)^{1-\eta} = z^B \left(k_t^B\right)^{\alpha(1-\eta)}, \tag{30}$$

$$z^B = \frac{(1-\alpha)\left[(1+\beta\varepsilon)\cdot \ell_\star^B(\theta) - 1\right](1-x)}{\varepsilon(1+\beta)\left(1+\ell_\star^B(\theta)\right)\Psi\left(1-\ell_\star^B(\theta)\right)^{\varepsilon}\left(\rho^B\right)^{\eta}}. \tag{31}$$

In the public regime, the effects of fiscal policy crucially depend on how second-period tax rates modify the intersectoral allocation of time between studying and working for young agents. These *time-reallocation effects* can be described by considering $\ell_\star^B$ as a function of θ, which yields the following results

Proposition 3 *Work time $\ell_\star^B$ depends on θ with the following properties:*

i. $\frac{\partial}{\partial\theta}\ell_\star^B(\theta) > 0$, $\lim_{\theta\to 1}\ell_\star^B(\theta) = 1$, $\lim_{\theta\to-\infty}\ell_\star^B(\theta) = \ell_{\min}^B$;

ii. there exists a unique $\bar{\theta}$ such that $\ell_\star^B\left(\bar{\theta}\right) = \ell_\star^A$;

iii. $\bar{\theta} > 0$;

iv. $\ell_\star^B(0) < \ell_\star^A$.

Property (i) is intuitive: lowering second-period tax rates induces young agents to study more and devote less time to work; heavily taxing adults forces individuals to work in the first period of life, in order to accumulate savings and rely on capital income in the second period. Symmetrically, subsidising adult generations reduces work time, bringing $\ell_\star^B$ towards the lower bound. Properties (ii)-(iii) define a critical tax rate: setting $\theta_t = \bar{\theta}$ in each period, work time is the same in the two regimes. We will refer to $\bar{\theta}$ as the *labour-neutral* tax rate, which is strictly positive by (iii). Property (iv) establishes that setting $\theta_t = 0$ implies lower work time in the public school regime.[7]

From (17) and (30), the physical-human capital ratio and the output-human capital ratio in both regimes converge to

$$\lim_{t\to\infty} k_t^i = k_{ss}^i = \left(z^i\right)^{\frac{1}{1-\alpha(1-\eta)}}, \qquad \lim_{t\to\infty} y_t^i = y_{ss}^i = \left(z^i\right)^{\frac{\alpha}{1-\alpha(1-\eta)}}, \tag{32}$$

long-run growth is determined by the learning technology

$$\lim_{t\to\infty} \left(Y_{t+1}^i / Y_t^i\right) = \varphi_{ss}^i = \Psi\left(1-\ell_\star^i\right)^{\varepsilon} \left(\rho^i y_{ss}^i\right)^{\eta}, \tag{33}$$

and accumulation rates equal, by (10) and (22),

$$z^A = (\alpha/\varepsilon\Psi)\left(\rho^A\right)^{-\eta}\left(1-\ell_\star^A\right)^{1-\varepsilon}, \tag{34}$$

$$z^B = (\alpha/\varepsilon\Psi)\left(\rho^B\right)^{-\eta}\left(1-\ell_\star^B(\theta)\right)^{1-\varepsilon}\left(\frac{1-x}{1-\theta}\right). \tag{35}$$

Equation (35) shows that fiscal policy effectiveness depends on the intertemporal allocation of education costs, which is determined in the public regime by the *tax ratio* $(1-x)(1-\theta)^{-1}$. Therefore, fiscal policy may affect growth and welfare through two channels: on the one hand, study time can be increased by reducing second-period tax rates; on the other hand, first-period tax rates and public spending may increase long-run growth by raising the accumulation rate. This dichotomy is crucial for results presented in the next section.

As regards welfare, individual lifetime utility in period t is the sum of three components (see Appendix):

$$U_t^i = \Lambda^i + (1+\alpha\beta)\log k_{t+1}^i + (1+\beta)\sum_{j=0}^{t}\log\varphi_j^i, \quad i = A, B. \tag{36}$$

The *static term* Λ depends on initial endowments and work time. The *accumulation term* varies only in the short run and converges to $(1+\alpha\beta)\log k_{ss}^i$. The last term in (36), instead, grows indefinitely, implying that individual

[7] Work time in the public system does not depend on the marginal effect of expenditures on learning (coefficient p^A depends on η, whereas p^B does not - see Appendix). This implies that work time differs in the two regimes when $\theta = 0$.

welfare exhibits a positive time-trend over generations: for t and t_0 large enough with $t > t_0$, the *growth term* can be rewritten as

$$(1+\beta)\sum_{j=0}^{t}\log\varphi_j^i \approx (1+\beta)\sum_{j=0}^{t_0}\log\varphi_j^i + (t-t_0)(1+\beta)\log\varphi_{ss}^i. \tag{37}$$

Expression (37) shows that the effects of knowledge transmission dominate, in terms of welfare levels, the static term Λ in the long run.

3 Tax policy analysis

This section studies how the relation between taxation and growth is influenced by time-reallocation and crowding-in effects. The analysis is positive in spirit, and describes the implications for growth and welfare of alternative tax policies. A policy is defined as a sequence $\{\rho_t^B, x_t, \theta_t\}$ implemented by fiscal authorities over the whole time-horizon $t = 0, ..., \infty$. Since there are three policy instruments, it is convenient to restrict the analysis to a subset of issues related to tax policy, focusing on the growth effects of alternative ways to finance public investment. Specifically, we fix instrument ρ^B by assuming that the public propensity to spend is maintained at the efficient level, and study the growth scenarios implied by different combinations of tax rates. This procedure allows us to describe the implications of re-distributing the burden of education across the individual life-cycle according to different criteria. Since any value of θ corresponds to a unique feasible level of work time, ρ^B and x are associated to each possible θ as follows:

$$\rho^B(\theta) = \eta\varepsilon^{-1}(1-\alpha)\left(1-\ell_\star^B(\theta)\right), \tag{38}$$

$$x(\theta) = \frac{\eta\varepsilon^{-1}\left(1-\ell_\star^B(\theta)\right)-\theta}{\ell_\star^B(\theta)}. \tag{39}$$

Equation (38) is formally analogous to the optimality condition obtained for the private propensity and represents the efficient public propensity to spend (section 3.1 shows that (38) indeed characterises Ramsey-optimal allocations). Expression (39) defines the unique value of x consistent with the budget constraint. The next section describes the Ramsey-optimal policy, which implies a high tax rate for young individuals. Section 3.2 investigates the consequences of alternative policies that shift the burden from the first to the second period of life.

3.1 Ramsey-optimal policy

The Ramsey-optimal policy is defined according to the standard criterion: a hypothetical central planner seeks the sequence of consumption levels, work time and school quality $\{c_t^*, d_t^*, \ell_t^*, E_t^*\}$ that maximises the discounted sum of lifetime utilities[8]

$$\Upsilon = \sum_{t=0}^{\infty} n\Phi^t \left(\log c_t + \beta \log d_{t+1}\right), \tag{40}$$

where $\Phi \in (0, 1)$ is the social discount factor, subject to the transition law of human knowledge (1)-(2), and to the aggregate resource constraint of the economy

$$K_{t+1} = K_t^{\alpha} \left[h_t \left(1 + \ell_t\right)\right]^{1-\alpha} - nc_t - nd_t - E_t h_t, \tag{41}$$

taking initial endowments (h_0, K_0) as given. The solution of the social problem determines the Ramsey-optimal allocation, and a Ramsey-optimal policy is a sequence of fiscal instruments $\{\rho_t^*, x_t^*, \theta_t^*\}$ that implements such an allocation. This requires to set tax rates and the public spending ratio $g_t \equiv E_t^B$ in order to satisfy the government budget constraint and the optimality conditions of the centralised problem. Formally, the following relations must hold in each period (see Appendix)

$$\rho_t^B = \frac{\eta}{\varepsilon} \left(1 - \alpha\right) \left(1 - \ell_t^B\right), \tag{42}$$

$$\rho_t^B = \left(1 - \alpha\right) \left(\theta_t + x_t \ell_t^B\right), \tag{43}$$

$$\frac{1 - \theta_{t+1}}{1 - x_t} = 1 + \frac{1}{\varepsilon} \left[1 - \eta - \left(1 - \varepsilon - \eta\right) \cdot \ell_{t+1}^B\right]. \tag{44}$$

Equation (42) is the efficient public propensity to spend in education, which indeed verifies (38); equation (43) is the government budget constraint; imposing the equality between market factor prices (w_t^B, R_t^B) and optimal marginal productivities yields equation (44).

There is a unique Ramsey-optimal policy compatible with convergence towards balanced growth, which features constant tax rates. The constant tax rate θ^* satisfying conditions (42)-(44) is recursively determined by the

[8] Note that we are not assuming *ex-ante* that ρ^B is kept at an efficient level: in this subsection, equation (38) is derived from the maximisation process.

system

$$\begin{aligned} \frac{1-\theta^*}{1-x^*} &= \phi(\theta^*), & (45) \\ x^* &= x(\theta^*), & (46) \\ \rho^* &= \rho^B(\theta^*), & (47) \end{aligned}$$

where $x(\theta^*)$ and $\rho^B(\theta^*)$ are defined by (38)-(39), and[9]

$$\phi(\theta) = \frac{1}{\varepsilon}\left[\varepsilon\left(1+\ell_\star^B(\theta)\right) + (1-\eta)\left(1-\ell_\star^B(\theta)\right)\right] > 1. \quad (48)$$

It derives from $\phi > 1$ that individuals are taxed more heavily in their first period of life: for example, when the learning technology exhibits constant returns to scale, $\varepsilon + \eta = 1$, Ramsey-optimal taxation implies $\phi(\theta^*) = 2$. More generally, with non-increasing returns to scale ($\varepsilon + \eta \leq 1$) the following results hold.

Proposition 4 *Under the Ramsey-optimal policy $\{\rho^*, x^*, \theta^*\}$, work time is lower in the public regime, and public propensity to spend in education is higher than private propensity:*

$$\begin{aligned} \ell_\star^B(\theta^*) &< \ell_\star^A, & (49) \\ \rho^* &> \rho^A. & (50) \end{aligned}$$

Proposition 4 can be interpreted as follows. In the private regime, individual study time and expenditures in education are below Ramsey-optimal levels, because finitely-lived selfish agents do not fully internalise the benefits of knowledge transmission. Ramsey-optimal policies cure this market incompleteness by increasing study time and the propensity to spend in education. However, the associated tax burden falls heavily on young workers ($\phi > 1$), whereas adults are generally subsidised.[10] High tax rates on young agents, low work time, and high propensity to spend in education drive down the long-run saving rate, which is generally higher in the private system:

---

[9] Equation (48) can be rewritten as $\phi = 1 + \left[\ell_\star^B(\theta^*) + \left(\frac{1-\eta}{\varepsilon}\right)\left(1-\ell_\star^B(\theta^*)\right)\right]$, where the term in square brackets is strictly positive, implying $\phi > 1$.

[10] It can be shown that $\theta^* < 0$ obtains when $\beta\eta < 0.5$ in the general case $\varepsilon + \eta \leq 1$. Moreover, $\beta\eta < 0.5$ is not strictly necessary to have $\theta^* < 0$: for example, when $\varepsilon + \eta = 1$ it follows from $\phi = 2$ that adults are subsidised whenever $x^* < 0.5$, which is always the case in our simulations.

Lemma 5 *Under the Ramsey-optimal policy $\{\rho^*, x^*, \theta^*\}$, the accumulation rate z^B is strictly lower than z^A if either $\varepsilon + \eta = 1$, or $\varepsilon + \eta < 1$ with a sufficiently low work time gap $\ell_\star^A - \ell_\star^B(\theta^*)$.*

Proof. Lemma 5 is proved as follows: from (34), (35), and (45),

$$\frac{z^A}{z^B(\theta^*)} = \phi(\theta^*) \cdot \left(\frac{1-\ell_\star^A}{1-\ell_\star^B(\theta^*)} \right)^{1-\varepsilon-\eta}. \tag{51}$$

If $\varepsilon + \eta = 1$ the term in round-brackets equals unity, and $\phi = 2$ implies $z^A > z^B(\theta^*)$. If $\varepsilon + \eta < 1$, expression (51) yields $z^A > z^B(\theta^*)$ provided the term in round-brackets (below unity) is more than offset by $\phi > 1$. These results imply that Ramsey-optimal policies generally involve crowding-out of physical capital, which influences knowledge formation via public spending. In terms of growth rates, the negative effect on savings can be offset by the benefits of human capital formation: the asymptotic growth ratio equals[11]

$$\left(\varphi_{ss}^A / \varphi_{ss}^B\right) = \phi(\theta^*)^{\frac{\alpha\eta}{1-\alpha(1-\eta)}} \left[\frac{1-\ell_\star^A}{1-\ell_\star^B(\theta^*)} \right]^{\varepsilon+\eta+\frac{\alpha\eta(1-\varepsilon-\eta)}{1-\alpha(1-\eta)}}, \tag{52}$$

where the term in square brackets is always below unity. Whether the whole expression is below unity depends on the parameters, and simulations suggest that the relative share α plays a crucial role in this regard. Low values of α tend to reduce the negative impact of crowding-out effects on $\varphi_{ss}^B(\theta^*)$ because the relative importance of K in production is limited. Conversely, high values of α strengthen crowding-out effects, and Ramsey-optimal taxation is more likely to be growth-reducing with respect to a pure private system. For example, setting $\varepsilon = \eta = 0.5$ and $\beta = 0.8$ implies $\varphi_{ss}^A/\varphi_{ss}^B = 0.946$ for $\alpha = 0.4$, whereas setting $\alpha = 0.6$ *ceteris paribus* yields $\varphi_{ss}^A/\varphi_{ss}^B = 1.052$.

The implications for welfare are as follows. As shown in the previous section, individual utility levels are crucially determined by the 'growth term' in the long run. Consequently, the welfare gap between the two regimes, $U^A - U^B$, reflects the sign of the growth gap $\varphi_{ss}^A - \varphi_{ss}^B$ once the economy has approached balanced growth. Figure 1 shows that for $\alpha = 0.4$, utility U^B under Ramsey-optimal taxation is always higher than U^A. When $\alpha = 0.6$,

[11] Substituting (32) in (33) yields $\varphi_{ss}^i = \Psi\left(1-\ell_\star^i\right)^\varepsilon \left(\rho^i\right)^\eta \left(z^i\right)^{\frac{\alpha\eta}{1-\alpha(1-\eta)}}$. Substituting (16) and (34) in this expression gives φ_{ss}^A, while substituting $\rho^* = \rho^B(\theta^*)$ and (35) yields φ_{ss}^B. Taking the ratio and substituting (45) gives (52).

on the other hand, $U^B > U^A$ only for a finite number of periods: in this case, long-run growth rates are such that $\varphi_{ss}^A > \varphi_{ss}^B(\theta^*)$, and individuals enjoy higher utility from the private system in the long run.

3.2 Alternative tax policies

This section studies the effects of alternative policies that shift the tax burden from the first to the second period of life. Firstly, we analyse the properties of *labour-neutral policies*: when tax rates are adjusted so that work time is equal between public and private regimes, economic activity in the public system is sustained by increased savings through a crowding-in mechanism, and long-run growth is unambiguously higher than under the private system regardless of parameter values. This implies that labour-neutral taxation yields higher growth and long-run welfare than under the Ramsey-optimal policy when the latter strategy brings intensive crowding-out. Secondly, we show that there generally exists a set of non-optimal tax rates yielding higher growth and long-run welfare with respect to the Ramsey policy, even when the latter strategy is growth-improving with respect to the private regime.

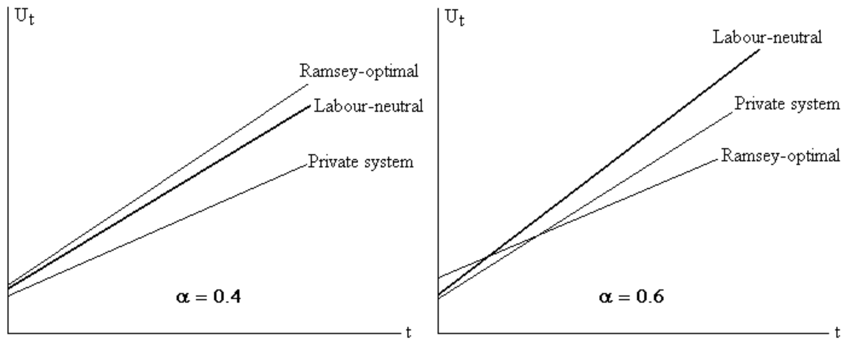
Labour-neutral taxation. Assume that the government implements labour-neutral taxation, as defined in Proposition 3. Setting $\theta_t = \bar{\theta}$ in each period, equations (38)-(39) define a constant propensity $\bar{\rho} = \rho^B\left(\bar{\theta}\right)$, and a constant tax rate $\bar{x} = x\left(\bar{\theta}\right)$ on young generations. The properties of the labour-neutral policy $\{\bar{\rho}, \bar{x}, \bar{\theta}\}$ are summarised in the following

Proposition 6 *Under the labour-neutral policy $\{\bar{\rho}, \bar{x}, \bar{\theta}\}$, young generations are subsidised ($\bar{x} < 0$), the accumulation rate is higher in the public regime ($z^B > z^A$), and public education guarantees higher growth and welfare at least in the long run ($\varphi_{ss}^B > \varphi_{ss}^A$).*

When labour supply effects are neutralised by tax policy, public education guarantees higher growth. The reason is that $\bar{x} < 0$ implies higher disposable income for young generations, and higher savings: this *crowding-in* effect sustains economic activity by raising the accumulation rate, which influences long-run growth through public spending:

$$z \uparrow \quad \Rightarrow \quad y_{ss} \uparrow \quad \Rightarrow \quad \rho y_{ss} \equiv g_\infty \equiv E_\infty \uparrow .$$

The major point to be emphasised here is that $\varphi_{ss}^B\left(\bar{\theta}\right)$ unambiguously exceeds φ_{ss}^A; that is, crowding-in effects yield higher long-run growth with respect



Graph (a)

t	Private System	Labour-Neutral	Ramsey-Optimal
0	9.694	9.696	10.084
1	11.526	11.639	12.021
2	13.374	13.592	13.974
3	15.226	15.548	15.932
4	17.079	17.504	17.890
5	18.932	19.460	19.849
6	20.785	21.416	21.808
7	22.638	23.372	23.767
8	24.491	25.328	25.725
9	26.344	27.284	27.684
10	28.196	29.240	29.643
11	30.049	31.196	31.602
12	31.902	33.152	33.561
13	33.755	35.108	35.519
14	35.608	37.064	37.478
15	37.461	39.020	39.437
16	39.314	40.976	41.396
17	41.167	42.932	43.355
18	43.020	44.888	45.313
19	44.873	46.844	47.272
20	46.726	48.800	49.231

Graph (b)

t	Private System	Labour-Neutral	Ramsey-Optimal
0	11.545	11.569	12.935
1	13.718	13.832	15.083
2	15.673	15.865	16.948
3	17.588	17.856	18.769
4	19.494	19.835	20.581
5	21.397	21.812	22.391
6	23.299	23.787	24.199
7	25.201	25.762	26.007
8	27.102	27.737	27.815
9	29.004	**29.712**	**29.623**
10	30.906	31.687	31.432
11	32.807	33.662	33.240
12	34.709	35.637	35.048
13	36.611	37.612	36.856
14	38.512	39.587	38.664
15	40.414	41.562	40.472
16	**42.316**	43.537	**42.280**
17	44.217	45.512	44.088
18	46.119	47.487	45.896
19	48.021	49.462	47.705
20	49.923	51.437	49.513

Figure 1: Time paths of lifetime utility under different tax policies with $\beta = 0.8$, $\varepsilon = \eta = 0.5$, $K_0 = 10$, $h_0 = 1$. Graph (a) assumes $\alpha = 0.4$: the Ramsey-optimal policy improves growth and welfare with respect to both the private system and the labour-neutral policy (see Figure 3 for details). Graph (b) assumes $\alpha = 0.6$: labour-neutral taxation yields higher long-run growth with respect to both the private system and the Ramsey-optimal policy; setting $\Psi = 230$ yields $\varphi_{ss}^{B}\left(\bar{\theta}\right) = 9.72\%$, $\varphi_{ss}^{A} = 5.65\%$, and $\varphi_{ss}^{B}\left(\theta^{*}\right) = 0.45\%$.

to the private system regardless of parameter values. As shown in section 3.1, Ramsey-optimal policies do not have this property, as the sign of the growth gap $\varphi_{ss}^{A} - \varphi_{ss}^{B}$ crucially depends on the value of α. Consequently, labour-neutral taxation may yield higher growth and long-run welfare than Ramsey-optimal taxation: when parameters are such that $\varphi_{ss}^{B}(\theta^{*})$ is lower than φ_{ss}^{A}, Proposition 6 implies $\varphi_{ss}^{B}\left(\bar{\theta}\right) > \varphi_{ss}^{A} > \varphi_{ss}^{B}(\theta^{*})$. Put differently, high values of α strengthen not only crowding-out, but also crowding-in effects, so that shifting the burden onto adult workers enhances growth when the share of physical capital in production is relatively high. This is confirmed by the numerical example described in Figure 1: consistent with Proposition 6, the labour-neutral policy is welfare-improving with respect to the private system regardless of the values of α. When $\alpha = 0.4$, Ramsey-optimal taxation yields higher utility levels than labour-neutral taxation, but the opposite result holds setting $\alpha = 0.6$: from period $t = 9$ onwards, labour-neutral taxation implies higher utility than the Ramsey-optimal policy.

It is important to note that Ramsey-optimal and labour-neutral policies cannot be Pareto-ranked[12]. By construction, any policy that shifts the burden of first-period education onto the second period of life implies an *income-redistribution effect* that brings welfare losses for adult agents in period zero. This is a typical 'first-father problem': public education of those who are young in $t = 0$ must be financed by a generation which does not receive any benefit from the newly-established school system.[13] Bearing this in mind, the policy implications of Figure 1 are nonetheless interesting: first, when crowding-out effects are important, Ramsey-optimal policies do not preserve the welfare of all generations with respect to a pure private system, whereas

[12] Figure 1 compares different policies in terms of lifetime utility levels. Alternatively, one might consider present-value streams of utilities yielded by the same policies, using a predetermined social discount factor. On the one hand, it is possible to construct ranges of values for Φ such that welfare gains from non-Ramsey policies might compensate, in present-value terms, the loss of intertemporal efficiency over some chosen time interval (t_0, t_1), abstracting from the first-father problem. On the other hand, present-value comparisons would not allow assessing the intergenerational distribution of benefits - which is a major focus of the analysis - since "social welfare" would in this case be thought of as a discounted sum over generations.

[13] The time-path of individual welfare depicted in Figure 1 refers to lifetime utility $U_t = U(c_t, e_{t+1})$, so that $U_0 = U(c_0, e_1)$ pertains to those who are young in period zero. Instead, second-period utility of the *first generation of fathers* - *i.e.* those adult at time zero - equals $\beta \log e_0$ and is necessarily higher under Ramsey-optimal taxation with respect to labour-neutral taxation.

labour-neutral policies always succeed in this regard. Second, knowledge transmission amplifies the effectiveness of policies that increase disposable income of savers, with possible gains in terms of growth and long-run welfare with respect to Ramsey-optimal taxation.

The Growth Curve. The labour-neutral policy described above is a peculiar tax rule which implements the same work time in a public system as in the private regime. When comparing the effects of alternative strategies for financing public education, this tax rule represents a useful device for obtaining analytical results on crowding-in effects. However, the role of crowding-in effects can be assessed in more general terms by considering the whole set of tax rates that shift the education burden onto adults. That is, crowding-in effects may be exploited by policies which are not necessarily labour-neutral, in order to preserve individual welfare of late-in-time generations. This point can be addressed by studying the relation between asymptotic growth and the tax rate on adults. Substituting (38) and (39) in (35), the asymptotic growth rate can be expressed as

$$\varphi_{ss}(\theta) = \xi \left(1 - \ell_{\star}^{B}(\theta)\right)^{\varepsilon+\eta+\frac{(1-\varepsilon-\eta)\alpha\eta}{1-\alpha(1-\eta)}} \left(\frac{1-x(\theta)}{1-\theta}\right)^{\frac{\alpha\eta}{1-\alpha(1-\eta)}}, \tag{53}$$

where ξ is a constant parameter. Expression (53) describes the effects of taxation on growth in terms of two factors, the level of study time and the *tax ratio* $(1-x(\theta))(1-\theta)^{-1}$. These terms react in opposite directions as the tax rate on adults varies: an increase in θ implies reduced study time, but also an increase in the tax ratio.[14] Figure 2 shows that the growth curve (53) has an inverted-U shape, which is explained as follows. From property (i) of Proposition 3, $\lim_{\theta\to-\infty}\left(\partial\ell_{\star}^{B}(\theta)/\partial\theta\right) = 0$. Consequently, the net effect of a marginal increase $d\theta$ on long-run growth is positive for low values of θ, due to the increase in the tax ratio. However, for high values of θ, the negative effect on growth of reduced study time dominates at the margin. It is worth noting that this relation is a peculiar feature of our model, which results from the assumption of productive public expenditures.

The inverted-U shape of the growth curve implies that there exists a unique tax rate $\hat{\theta}$ which maximises long-run growth. Since $\hat{\theta}$ generally differs from the Ramsey-optimal rate θ^{*}, there exists a tax rate $\theta^{n} \neq \theta^{*}$ yielding

[14] Equation (39) implies that $x(\theta)$ is decreasing in θ. Therefore, a marginal increase in θ increases the tax ratio in (53).

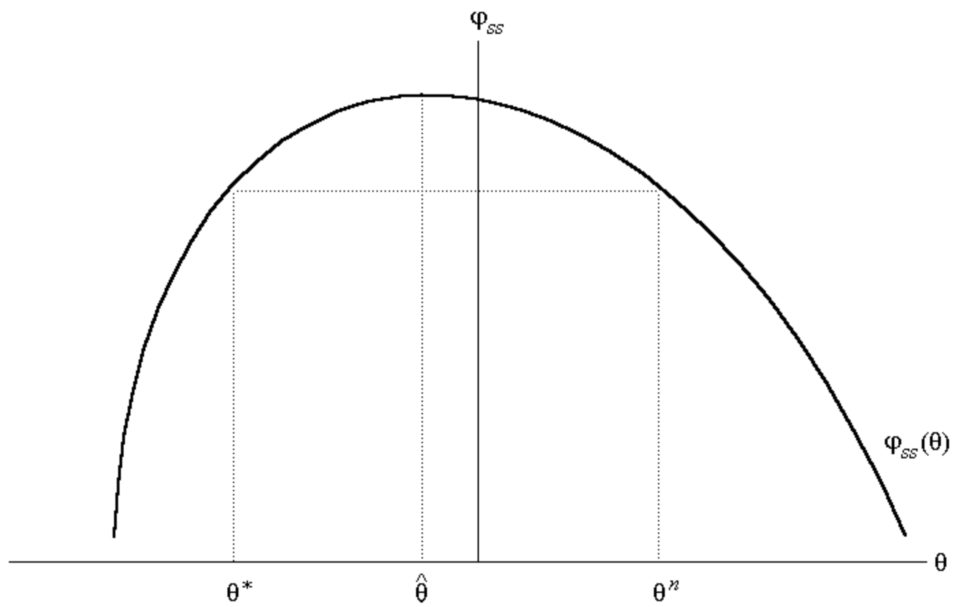


Figure 2: The growth curve (53) is an inverted-U relation between the asymptotic growth rate and the tax rate on adults. Setting $\alpha = 0.4$ with the same parameters of Figure 1, the Ramsey-optimal tax rate is $\theta^* = -21.9\%$ and the growth-equivalent rate is $\theta^n = 14.3\%$. Any value $\theta \in (\theta^*, \theta^n)$ is growth-improving with respect to the Ramsey-optimal policy.

a long-run growth rate equal to that obtained under the Ramsey-optimal policy (see Figure 2). We will refer to θ^n as the *growth-equivalent* tax rate on adults: using equations (38) and (39), a growth-equivalent policy $\{\rho^n, x^n, \theta^n\}$ is defined by a tax rate on adults $\theta^n \neq \theta^*$, an efficient propensity to spend in education $\rho^n = \rho^B(\theta^n)$, and a balanced tax rate $x^n = x(\theta^n)$, such that $\varphi^B_{ss}(\theta^n) = \varphi^B_{ss}(\theta^*)$. The properties of the growth-equivalent policy crucially depend on whether $\theta^* < \theta^n$:

Proposition 7 *When $\theta^* < \theta^n$, the accumulation rate and the output-human capital ratio are higher under the growth-equivalent policy if either $\varepsilon + \eta = 1$, or $\varepsilon + \eta < 1$ with a sufficiently low work time gap $\ell^B_\star(\theta^n) - \ell^B_\star(\theta^*)$:*

$$z^B(\theta^n) > z^B(\theta^*), \tag{54}$$
$$y^B_{ss}(\theta^n) > y^B_{ss}(\theta^*). \tag{55}$$

Figure 2 depicts the case considered in Proposition 7: the Ramsey-optimal tax rate θ^* lies to the left of the growth-maximising rate $\hat{\theta}$, implying $\theta^* < \hat{\theta} < \theta^n$. A little algebra shows that $\theta^* < \hat{\theta}$ requires[15]

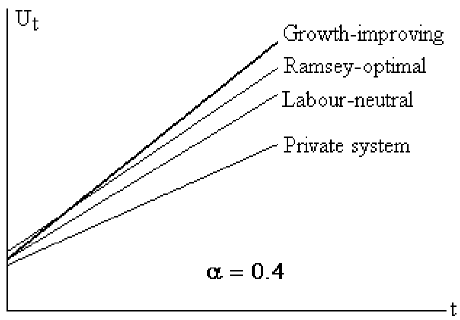
$$1/2 < \alpha\eta \left(1 - \alpha\varepsilon\right)^{-1} \left(1 - \ell_\star^B \left(\hat{\theta}\right)\right) \cdot \left(\tau'_\theta / \ell'_\theta\right). \tag{56}$$

Condition (56) is always satisfied in our simulations with varying parameters, but analytical proof of its general validity is quite difficult to obtain. The point is that when (56) is satisfied, there exists a *continuum* of second-period tax rates that redistribute the burden in favour of young generations and imply higher growth with respect to the Ramsey-optimal policy. This continuum is represented by the interval (θ^*, θ^n), as shown in Figure 2.

This result reinforces our previous conclusions about the importance of crowding-in effects: there generally exists a set of *growth-improving* tax rates $\tilde{\theta} \in (\theta^*, \theta^n)$ yielding higher long-run growth and welfare with respect to the Ramsey-optimal policy, and this happens even for relatively low values of α. Figure 3 considers the same parameter values used in Figure 1, with $\alpha = 0.4$. The Ramsey-optimal tax rate is $\theta^* = -21.9\%$ and the associated growth-equivalent tax rate is $\theta^n = 14.3\%$. In this case, labour-neutral taxation is not growth-improving with respect to the Ramsey-optimal policy (as shown in Figure 1) because $\bar{\theta} = 14.9\% > \theta^n$. Choosing a growth-improving tax rate $\tilde{\theta} \in (\theta^*, \theta^n)$, fiscal authorities may redistribute the burden in favour of young generations and obtain $\varphi_{ss}^B\left(\tilde{\theta}\right) > \varphi_{ss}^B\left(\theta^*\right)$. In the numerical example of Figure 3 we set $\theta = 4\%$, which implies higher growth in the long run: consequently, individual utility levels are above those obtained under the Ramsey-optimal policy from period $t = 8$ onwards.

It is worth noting that these results recall the logic of *Gale-type intergenerational transfers*. Gale (1973) showed that, for a two-generations pure exchange economy, the first generation can raise future welfare by renouncing part of its claim to the endowments benefitting the second generation, which in turn transmits a claim to its successor, and so on. In the present context, the income-redistribution effect amounts to the share of claims on human capital not received by adults at time zero, and the impact of such policies

[15] For simplicity, (56) is derived with constant returns to scale $\varepsilon + \eta = 1$. The left-hand side of (56) is the *Ramsey-optimal tax ratio*, which equals $1/\phi\left(\theta^*\right) = 0.5$ by (45). The right-hand side is the *growth-maximising tax ratio*, where τ'_θ is the total derivative of the tax ratio with respect to θ, and $\ell'_\theta = \partial \ell_\star^B\left(\theta\right)/\partial\theta$, with τ'_θ and ℓ'_θ both evaluated at $\theta = \hat{\theta}$.



	Private	Lab-Neutral	Growth-Imp.	Ramsey-Opt
ℓ	0.902	0.902	0.893	0.876
x	-	-5.6%	7.4%	39.0%
θ	-	14.9%	4.0%	-21.9%
ρ	5.9%	5.9%	6.4%	7.4%
z	0.032	0.040	0.031	0.016
y_{ss}	0.180	0.200	0.176	0.127
φ_{ss}	1.029	1.087	1.105	1.088

Figure 3: Time paths of lifetime utility under different tax policies with $\beta = 0.8$, $\varepsilon = \eta = 0.5$, $K_0 = 10$, $h_0 = 1$, and $\Psi = 230$. Values for the private system, Ramsey-optimal and labour-neutral policies are the same as in Figure 1 - graph (a), now compared with the growth-improving policy $\tilde{\theta} = 4\%$.

on long-run growth and welfare depends on the interplay between physical capital accumulation, time-reallocation effects, and knowledge transmission. From a policymaking perspective, the rationale for such policies would hinge on the possibility that the government is able to sterilise the welfare loss for adults at time zero. In this regard, the assumption of a balanced public budget is crucial, since the first-father problem arises as long as first-period education at time zero *must* be financed at time zero. This suggests a powerful role for the use of public debt: for example, the government may run positive debt at time zero to finance initial spending, and then smooth service repayments over time according to a calibrated fiscal rule. Whether debt policy rules for intergenerational *fiscal* fairness are compatible with socially optimal growth rates may be an interesting topic for future research.

4 Conclusions

This paper analysed the effects of alternative tax policies on growth and welfare in a model with knowledge transmission and labour-enhancing public investment. Fiscal policy affects growth and welfare through two channels: the relation between learning and public expenditures, and time-reallocation effects induced by variations in prospective tax rates. These two mechanisms imply an inverted-U relation between long-run growth and second-period tax rates, so that a generic growth rate can be obtained by means of opposite financing strategies: 'taxing adults' or 'taxing the young'. Under Ramsey-optimal policies, the education burden falls on young workers, reducing the saving rate and crowding out physical capital. Consequently, Ramsey-optimal strategies for financing public education may reduce growth and long-run welfare with respect to a pure private system, depending on the relative importance of physical capital in production. Alternative policies that shift the burden onto adults instead protect the welfare of all generations. On the one hand, there always exists a labour-neutral policy under which growth and welfare are unambiguously higher than under the private system, by virtue of the crowding-in mechanism. On the other hand, there generally exists a continuum of non-optimal tax rates under which long-run growth and welfare are higher than with the Ramsey-optimal policy.

More generally, this paper studied how the relation between taxation and growth is influenced by the assumption of labour-enhancing educational expenditures: most previous literature emphasised the time-reallocation effects induced by distortionary income taxation, using models where crowding-in mechanisms play a minor role. But crowding-in effects are important when educational expenditures affect learning, because knowledge transmission amplifies the effectiveness of policies that increase the disposable income of savers. In particular, if the first generation renounces part of its claims on human capital, crowding-in effects sustain economic growth and welfare in the long run.

Appendix

Proof of Lemma 1 and Lemma 2. Substituting the first-order condition (10) in the accumulation rule (13), the dynamics of ℓ_t^A are described by

$$\ell_{t+1}^A = \varepsilon q^A \frac{\ell_t^A}{1-\ell_t^A} - p^A,$$

where $q^A = \beta\left(1-\alpha\right)\left(\alpha+\alpha\beta\right)^{-1}$, and $p^A = 1 + \left(1+\beta\eta\right)\left(1-\alpha\right)\left(\alpha+\alpha\beta\right)^{-1}$. This expression is analogous to (25): since q^B and p^B are constant when $\theta_t = \theta$ in each period, Lemma 1 and Lemma 2 can be proved by studying the generic dynamic equation

$$\ell_{t+1} = \varepsilon q \frac{\ell_t}{1-\ell_t} - p. \tag{A1}$$

Taking the limits on the right hand side of (A1) we obtain $\lim_{\ell_t \to 0} \ell_{t+1} = -p$ and $\lim_{\ell_t \to 1} \ell_{t+1} = +\infty$, that imply the existence of a stationary solution $\ell_{t+1} = \ell_t$. Rewriting (A1) as $\ell_{t+1} = -\left(\varphi_{\ell_t}/\varphi_t\right) q\ell_t - p$, the derivative of the right hand side with respect to ℓ_t is

$$\chi\left(\ell_t\right) = \frac{\partial \ell_{t+1}}{\partial \ell_t} = -q\left[\frac{\varphi_{\ell_t}}{\varphi_t} + \ell_t\left(\frac{\varphi_{\ell_t\ell_t}}{\varphi_t}\right) - \ell_t\left(\frac{\varphi_{\ell_t}}{\varphi_t}\right)^2\right] > 0,$$

implying that the stationary equilibrium $\ell_\star$ is unique. Setting $\ell_{t+1} = \ell_t$ in (A1) gives a second-order equation in ℓ with two roots of opposite sign: since $p > 1$, the positive root is $\ell_\star$ as defined by equations (14) and (28). Evaluating $\chi\left(\ell_t\right)$ at $\ell_t = \ell_\star$ gives

$$\chi\left(\ell_\star\right) = \frac{p+\ell_\star}{\ell_\star} - q\ell_\star\left[\frac{\varphi_{\ell_t\ell_t}}{\varphi_t} - \left(\frac{\varphi_{\ell_t}}{\varphi_t}\right)^2\right] > 1 + \frac{p}{\ell_\star} > 1,$$

hence work time displays unstable dynamics outside the stationary equilibrium $\ell_\star$. Consequently, work time jumps at the optimal level $\ell_\star$ in period zero and is constant thereafter.

Proof of Proposition 3. Since q^B and p^B depend on the tax rate on adult generations, optimal work time in the public regime is a function of θ. To simplify notation, we define $\Gamma = 1-\theta$ and study the function $\ell_\star^B(\Gamma)$. Setting $\Omega = 1 - \varepsilon q^B - p^B$ we can write, by (28),

$$\ell_\star^B = (1/2)\left[\Omega + \sqrt{(\Omega)^2 + 4p^B}\right], \tag{A2}$$

which implies

$$\frac{\partial \ell_\star^B}{\partial \Gamma} = \frac{1}{2}\left(\Omega_\Gamma + \frac{2\Omega_\Gamma\Omega + 4p_\Gamma^B}{2\sqrt{\Omega^2 + 4p^B}}\right), \tag{A3}$$

where $p_\Gamma^B = \partial p^B/\partial\Gamma > 0$ and $\Omega_\Gamma = \partial\Omega/\partial\Gamma < 0$. We firstly prove that $\ell_\star^B$ decreases with Γ. The proof is by contradiction: assume that $\partial\ell_\star^B/\partial\Gamma > 0$: by (A3), this requires (recalling that $\Omega_\Gamma < 0$)

$$\Omega + \sqrt{(\Omega)^2 + 4p^B} < -2\left(p_\Gamma^B/\Omega_\Gamma\right). \tag{A4}$$

Evaluating p_Γ^B and Ω_Γ on the basis of (26)-(27) gives

$$-\frac{p_\Gamma^B}{\Omega_\Gamma} = \frac{1}{1+\beta\varepsilon}. \tag{A5}$$

Substituting (A5) and $2\ell_\star^B = \Omega + \sqrt{(\Omega)^2 + 4p^B}$ in (A4) yields

$$\ell_\star^B < \frac{1}{1+\beta\varepsilon},$$

which is absurd because it violates (29). Therefore $\partial\ell_\star^B/\partial\Gamma < 0$, *i.e.* $\ell_\star^B$ increases with θ. From (26), (27) and (A2), we have $\lim_{\Gamma\to 0}\ell_\star^B = 1$. Taking the limit as $\Gamma \to \infty$ it can also be shown that $\lim_{\Gamma\to\infty}\ell_\star^B = \ell_{\min}^B$, which completes the proof of property (i). Property (ii) derives from property (i): since $\ell_\star^B$ decreases monotonically from 1 to $\ell_{\min}^B$ as Γ goes from 0 to $+\infty$, there is a unique intersection $\ell_\star^B = \ell_\star^A$ (the intersection exists because $\ell_\star^A > \ell_{\min}^A > \ell_{\min}^B$). As regards property (iv), when $\Gamma = 1$ (that is, when $\theta = 0$), we have $q^A \equiv q^B = q$, and optimal work time in the two economies differs only because $p^A > p^B$: if $\partial\ell_\star^i/\partial p^i > 0$, we can conclude that $\ell_\star^A > \ell_\star^B$. Evaluating $\partial\ell_\star^i/\partial p^i$ and substituting (A1) in the resulting expression gives

$$\left(1-\ell_\star^i\right)\left[\sqrt{\left(1-\varepsilon q - p^i\right)^2 + 4p^i}\right]^{-1} > 0,$$

hence $\ell_\star^B < \ell_\star^A$ when $\theta = 0$. Property (iii) is a corollary: since $\partial\ell_\star^B/\partial\theta > 0$, a strictly positive value $\theta = \bar{\theta}$ is required to obtain $\ell_\star^B = \ell_\star^A$.

Ramsey-optimal policy. The Ramsey problem is to maximise (40) subject to (1)-(2) and (41). Setting the Lagrangean

$$\begin{aligned}\Pi \;=\; & \sum_{t=0}^{\infty}\Big\{n\Phi^t\left(\log c_t + \beta\log d_{t+1}\right) + \lambda_t^h\left[h_t\Psi\left(1-\ell_t\right)^\varepsilon E_t^\eta - h_{t+1}\right] \\ & +\lambda_t^K\left[K_t^\alpha h_t^{1-\alpha}\left(1+\ell_t\right)^{1-\alpha} - nc_t - nd_t - E_t h_t - K_{t+1}\right]\Big\},\end{aligned}$$

where λ_t^h and λ_t^K are multipliers attached to human and physical capital transition laws respectively, optimality conditions for an interior solution read

$$\begin{aligned}
\lambda_{t+1}^K d_{t+1}^* &= \lambda_t^K \beta c_t^*, && \text{(A6)}\\
R_{t+1}^* \lambda_{t+1}^K &= \lambda_t^K, && \text{(A7)}\\
\lambda_t^h \varphi_t^* &= \lambda_{t-1}^h - \lambda_t^K \left[w_t^* \left(1+\ell_t^*\right) - E_t^* \right], && \text{(A8)}\\
\lambda_t^K w_t^* &= -\varphi_{\ell_t}^* \lambda_t^h, && \text{(A9)}\\
\lambda_t^K &= \varphi_{E_t}^* \lambda_t^h, && \text{(A10)}
\end{aligned}$$

where w_t^* and R_t^* indicate optimal marginal productivities. From (A8) and (42), the Ramsey-optimal allocation requires

$$\lambda_{t-1}^h = \lambda_t^K \left(1-\alpha\right) y_t \left[1+\ell_t - \left(\eta/\varepsilon\right)\left(1-\ell_t\right) - \left(\varphi_t/\varphi_{\ell_t}\right)\right],$$

which implies

$$\frac{\lambda_{t-1}^h H_t}{\lambda_{t-1}^K K_t} = \frac{1-\alpha}{\alpha}\left(1+\ell_t - \frac{\eta}{\varepsilon}\left(1-\ell_t\right) - \frac{\varphi_t}{\varphi_{\ell_t}}\right). \tag{A11}$$

Substituting (A9) in (22) we obtain

$$\frac{1-\theta_{t+1}}{1-x_t} = \frac{\alpha}{1-\alpha}\cdot\frac{\lambda_t^h H_{t+1}}{\lambda_t^K K_{t+1}}. \tag{A12}$$

Setting (A11) one period later and substituting (A12) gives condition (44) in the text.

Proof of Proposition 4. Substituting (48) in (45) gives

$$\theta^* = \frac{\phi\eta - \left[\phi\left(\varepsilon+\eta\right)-\varepsilon\right]\ell_\star^B}{\varepsilon\left(\ell_\star^B+\phi\right)}. \tag{A13}$$

We prove inequality (49) by showing that $\theta^* < \bar{\theta}$. The proof is by contradiction: supposing $\theta^* \geq \bar{\theta}$, Proposition 3 would imply

$$\ell_{\min}^A < \ell_\star^A \leq \ell_\star^B\left(\theta^*\right), \tag{A14}$$

and, since $\bar{\theta} > 0$, also θ^* must be strictly positive: by (A13), this requires

$$\ell_\star^B\left(\theta^*\right) \leq \frac{\phi\eta}{\phi\eta+\varepsilon\left(\phi-1\right)} \tag{A15}$$

Combining (A13) with (A15) and substituting $\ell^A_{\min}$ by (15) gives

$$\frac{1+\beta\eta}{1+\beta\varepsilon+\beta\eta}<\frac{\phi\eta}{\phi\eta+\varepsilon\left(\phi-1\right)},$$

which reduces to

$$\phi<1+\beta\eta. \tag{A16}$$

Substituting ϕ by (48) and rearranging terms gives

$$\ell^B_{\star}\left(\theta^{*}\right)>\frac{1-\eta-\varepsilon\eta\beta}{1-\eta-\varepsilon}, \tag{A17}$$

which is absurd because $\varepsilon>\varepsilon\eta\beta$ implies that the right-hand side of (A17) is greater than unity. This proves that $\theta^{*}<\bar{\theta}$, which implies $\ell^A_{\star}>\ell^B_{\star}\left(\theta^{*}\right)$. This in turn implies $\rho^{*}>\rho^A$ by virtue of equations (16) and (38).

Individual welfare. Substituting consumers' first-order conditions respectively in (8)-(9) and (19)-(20), optimal levels of first-period consumption are $c^A_t=\zeta^A Y_t$ and $c^B_t=\zeta^B Y_t$, where

$$\zeta^A = \frac{\delta\left(1-\alpha\right)}{1+\beta}\left(\frac{1-\eta-\left(1-\varepsilon\eta\right)\ell^A}{n\varepsilon\left(1+\delta\ell^A\right)}\right), \tag{A18}$$

$$\zeta^B = \frac{\delta\left(1-\alpha\right)}{\varepsilon\left(1+\beta\right)}\left[\frac{1-\ell^B_t\left(1-\varepsilon\right)}{1+\delta\ell}\right]\left(1-x\right), \tag{A19}$$

and, by (9) and (20), second-period consumption is

$$d^i_{t+1}=\alpha\beta\left(Y^i_t/K^i_{t+1}\right)\zeta^i Y^i_{t+1},\qquad i=A,B. \tag{A20}$$

Setting $K^i_{t+1}/Y^i_t=\tilde{z}^i$, equations (17) and (30) imply $\tilde{z}^i=z^i\Psi\left(1-\ell^i_{\star}\right)^{\varepsilon}\left(\rho^i\right)^{\eta}$. Using this result and substituting (A18)-(A19)-(A20) in (7), lifetime utility may be written as

$$U^i_t=\log\zeta^i+\beta\log\alpha\beta\left(\zeta^i/\tilde{z}^i\right)+\log Y^i_t+\beta\log\left(\tilde{z}^i Y^i_t\right)^{\alpha}\left(H^i_{t+1}\right)^{1-\alpha}, \tag{A21}$$

where the last term derives from $Y_{t+1}=K^{\alpha}_{t+1}H^{1-\alpha}_{t+1}$. Substituting, $\log Y^i_t=\log k^i_{t+1}+\log H^i_{t+1}-\log\tilde{z}^i$ in (A21) yields

$$U^i_t=\left(1+\beta\right)\log\left(\frac{\zeta^i}{\tilde{z}^i}\right)+\beta\log\alpha\beta+\left(1+\alpha\beta\right)\log k^i_{t+1}+\left(1+\beta\right)\log\left(H^i_{t+1}\right). \tag{A22}$$

From (3), $H^i_{t+1} = H^i_0 \prod_{j=0}^{t} \varphi^i_j$: substituting this expression in (A22) gives eq.(36) in the text, where static terms equal

$$\begin{aligned}
\Lambda^A &= \beta \log \alpha\beta (1-\alpha) + \log \left[\frac{1-\eta-\ell^A_\star (1-\varepsilon-\eta)}{\ell^A_\star (1+\beta\eta+\beta\varepsilon) - 1 - \beta\eta} \right]^{1+\beta} \left(H^A_0\right)^{1+\beta}, \\
\Lambda^B &= \beta \log \alpha\beta (1-\alpha) + \log \left[\frac{1-(1-\varepsilon)\ell^B_\star}{\ell^B_\star (1+\beta\varepsilon) - 1} \right]^{1+\beta} \left(H^B_0\right)^{1+\beta}.
\end{aligned}$$

Proof of Proposition 6. By Proposition 3, $\ell^A_\star = \ell^B_\star \left(\bar{\theta}\right) = \ell_\star$, and, by (16) and (38), $\rho^A = \rho^B \left(\bar{\theta}\right) = \bar{\rho}$. From (18) and (31), we obtain

$$\frac{z^A}{z^B} = \frac{\ell_\star (1+\beta\eta+\beta\varepsilon) - 1 - \beta\eta}{\left[\ell_\star (1+\beta\varepsilon) - 1\right] (1-\bar{x})}. \tag{A23}$$

Substituting (34) and (35) in the left-hand side of (A23) gives

$$\bar{\theta} = \eta\beta (1-\ell_\star) \left[\ell_\star (1+\beta\varepsilon) - 1\right]^{-1} > 0. \tag{A24}$$

We now prove that $\bar{x} < 0$. From (39), $\bar{x}\ell_\star = (1-\ell_\star)\frac{\eta}{\varepsilon} - \bar{\theta}$: substituting (A24) in this expression implies that $\bar{x} > 0$ only if

$$\ell_\star (1+\beta\varepsilon) > 1 + \beta\varepsilon, \tag{A25}$$

which is absurd since $\ell_\star < 1$. Therefore, young generations are subsidised. Since $\bar{\theta} > 0 > \bar{x}$, it derives from (34)-(35) that $z^B > z^A$. This implies, by (32), that $y^B_{ss} > y^A_{ss}$ and $k^B_{ss} > k^A_{ss}$. Consequently, school quality is asymptotically higher in the public regime ($\bar{\rho} y^B_{ss} > \bar{\rho} y^A_{ss}$), implying higher long-run growth:

$$\Psi (1-\ell_\star)^\varepsilon \left(\bar{\rho} y^A_{ss}\right)^\eta = \varphi^A_{ss} < \varphi^B_{ss} = \Psi (1-\ell_\star)^\varepsilon \left(\bar{\rho} y^B_{ss}\right)^\eta. \tag{A26}$$

As regards welfare, it follows from (37) that $\varphi^A_{ss} < \varphi^B_{ss}$ implies $U^A_t < U^B_t$ at least in the long run.

Proof of Proposition 7. Substituting (38) in (35) and taking the ratio between accumulation rates under growth-equivalent and Ramsey-optimal policies yields

$$\frac{z^B(\theta^n)}{z^B(\theta^*)} = \left(\frac{1-\ell^B_\star(\theta^n)}{1-\ell^B_\star(\theta^*)} \right)^{1-\varepsilon-\eta} \left[\frac{1-x(\theta^n)}{1-\theta^n} \cdot \frac{1-\theta^*}{1-x(\theta^*)} \right]. \tag{A27}$$

Given $\theta^* < \theta^n$, the term in square brackets in (A27) is always greater than unity because the tax ratio increases with θ. If $\varepsilon + \eta = 1$ the first term reduces to 1, implying $z^B(\theta^n) > z^B(\theta^*)$. If $\varepsilon + \eta < 1$, the whole expression exceeds unity provided the work time gap $\ell_\star^B(\theta^n) - \ell_\star^B(\theta^n)$ is sufficiently low. Analogous reasoning yields $y_{ss}^B(\theta^n) > y_{ss}^B(\theta^*)$ in the respective cases, because y_{ss}^B is an increasing function of z^B by (32).

References

Barro, R., Sala-i-Martin, X. (1995). 'Economic Growth'. New York: McGraw-Hill.

Benhabib, J., Spiegel, M.M. (1994). 'The role of human capital in economic development: evidence from aggregate cross-country data'. Journal of Monetary Economics 34 (2): 143-173.

Bouzahzah, M., de la Croix, D., Docquier, F. (2002). 'Policy reforms and growth in computable OLG economies'. Journal of Economic Dynamics and Control, 26 : 2093-2113.

Bovenberg, A.L., van Ewijk, C. (1997). 'Progressive taxes, equity, and human capital accumulation in an endogenous growth model with overlapping generations'. Journal of Public Economics, 64: 153-179.

Buiter, W., Kletzer, K. (1995). 'Capital mobility, fiscal policy, and growth, under self-financing of human capital formation'. Canadian Journal of Economics, 28 : S163-S194.

De Gregorio, J. (1996). 'Borrowing constraints, human capital accumulation, and growth'. Journal of Monetary Economics, 37: 49-71.

De Gregorio, J., Kim, S.-J. (2000). 'Credit markets with differences in abilities: education, distribution and growth'. International Economic Review, 41 (3): 579-607.

de la Croix, D., Michel, P. (2002). 'A theory of economic growth - Dynamics and policy in overlapping generations'. Cambridge: Cambridge University Press.

Docquier, F., Michel, P. (1999). 'Education subsidies, social security and growth: the implications of a demographic shock'. Scandinavian Journal of Economics, 101 (3): 425-440.

Eckstein, Z., Zilcha, I. (1994). 'The effects of compulsory schooling on growth, income distribution, and welfare'. Journal of Public Economics, 54: 339-359.

Gale, D. (1973), 'Pure exchange equilibrium in dynamic economic models'. Journal of Economic Theory, 6: 12-36.

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.

Heijdra, B., Ligthart, J. (2000). 'The dynamic macroeconomic effects of tax policy in an overlapping generations model'. Oxford Economic Papers, 52: 677-701.

Hendricks, L. (1999). 'Taxation and long-run growth'. Journal of Monetary Economics, 43: 411-434.

Lucas, R. (1988). 'On the mechanics of economic development'. Journal of Monetary Economics, 22: 3-42.

Meijdam, L. (1998). 'Taxes, growth and welfare in an endogenous growth model with overlapping generations'. Working paper, Department of Economics, Tilburg University.

Yakita, A. (2003). 'Taxation and growth with overlapping generations'. Journal of Public Economics, 87: 467-487.

Working Papers of the Institute of Economic Research

(PDF-files of the Working Papers can be downloaded at www.wif.ethz.ch/research.)

05/40 T.M. Steger and L. Bretschger
Globalization, the Volatility of Intermediate Goods Prices and Economic Growth

05/39 H. Egli
A New Approach to Pollution Modelling in Models of the Environmental Kuznets Curve

05/38 S. Valente
Genuine Dissaving and Optimal Growth

05/37 K. Pittel, J.-P. Amigues and T. Kuhn
Endogenous Growth and Recycling: A Material Balance Approach

05/36 L. Bretschger and K. Pittel
Innovative investments, natural resources, and intergenerational fairness: Are pension funds good for sustainable development?

04/35 T. Trimborn, K.-J. Koch and T.M. Steger
Multi-Dimensional Transitional Dynamics: A Simple Numerical Procedure

04/34 K. Pittel and D.T.G. Rübbelke
Private Provision of Public Goods: Incentives for Donations

04/33 H. Egli and T.M. Steger
A Simple Dynamic Model of the Environmental Kuznets Curve

04/32 L. Bretschger and T.M. Steger
The Dynamics of Economic Integration: Theory and Policy

04/31 H. Fehr-Duda, M. de Gennaro, R. Schubert,
Gender, Financial Risk, and Probability Weights

03/30 T.M. Steger
Economic Growth and Sectoral Change under Resource Reallocation Costs

03/29 L. Bretschger
Natural resource scarcity and long-run development: central mechanisms when conditions are seemingly unfavourable

03/28 H. Egli
The Environmental Kuznets Curve – Evidence from Time Series Data for Germany

03/27 L. Bretschger
Economics of technological change and the natural environment: how effective are innovations as a remedy for resource scarcity?

03/26 L. Bretschger, S. Smulders
Sustainability and substitution of exhaustible natural resources. How resource prices affect long-term R&D-investments

03/25 T.M. Steger
On the Mechanics of Economic Convergence

03/24 L. Bretschger
Growth in a Globalised Economy: The Effects of Capital Taxes and Tax Competition

02/23 M. Gysler, J.Kruse and R. Schubert
Ambiguity and Gender Differences in Financial Decision Making: An Experimental Examination of Competence and Confidence Effects

01/22 S. Rutz
Minimum Participation Rules and the Effectiveness of Multilateral Environmental Agremments

01/21 M. Gysler, M. Powell, R. Schubert
How to Predict Gender-Differences in Choice Under Risk: A Case for the Use of Formalized Models

00/20 S.Rutz, T. Borek
International Environmental Negotiation: Does Coalition Size Matter?

00/19 S. Dietz
Does an environmental Kuznets curve exist for biodiversity?