EUROPEAN UNIVERSITY INSTITUTE

DEPARTMENT OF ECONOMICS

EUI Working Paper ECO No. 99/14

Public Policies and Education, Economic Growth and Income Distribution

Günther Rehme

BADIA FIESOLANA, SAN DOMENICO (FI)

Printed in Italy in April 1999
European University Institute
Badia Fiesolana
I-50016 San Domenico (FI)
Italy

Public Policies and Education, Economic Growth and Income Distribution*

Günther Rehme†

Technische Universität Darmstadt

February 1999

*This paper is taken from Chapter 4 of my dissertation defended at the EUI in December 1998. I would like to thank Tony Atkinson, James Mirrlees, and Robert Waldmann for helpful advice and stimulating suggestions. I have also benefitted from comments by Frank Cowell, Jean-Yves Duclos, Lavan Mahadeva, Eugene Smolensky, and participants of the conferences SSCW98 in Vancouver, EEA98 in Berlin and IIPF98 in Córdoba, Argentina. Of course, all errors are my own. Financial support by the Deutscher Akademischer Austauschdienst (DAAD), grant no. 522-012-516-3, and an Erasmus grant for Nuffield College, Oxford, in 1995 are gratefully acknowledged.

†*Correspondence*: TU Darmstadt, FB 1/VWL 1, Schloss, D-64283 Darmstadt, Germany. phone: +49-6151-162219; fax: +49-6151-165553; E-mail: rehme@hrzpub.tu-darmstadt.de

Abstract

This paper provides an analysis relating economic growth, human capital composition, income distribution and public education policies. In the model human capital is 'lumpy' and only the high skilled carry it. The government chooses capital taxes to finance education, which directly affects growth, the *number* of high skilled people and wages. Growth and income equality depend positively on the productivity of the education sector. Analyzing various policies I show that e.g. the preferred policy of the unskilled is growth maximizing, whereas that of skilled labour leads to less education, lower growth and more wage inequality. The paper's public policy analysis provides an explanation for the recent observation of high growth and relatively low income inequality in some highly competitive economies.

KEYWORDS: Growth, Distribution, Human Capital, Public Education, Public Policy

JEL Classification: O41, I2, H2, D3, D6

1 Introduction

As markets become more integrated ('globalization'), human capital is assuming a pivotal role in policy debates (especially in OECD countries) on the maintenance of international competitiveness. The experience of some East Asian, high growth countries suggests that empirically there is a positive link from the provision of education to income equality and growth.[1] This paper offers a theoretical explanation of the stylized fact and contributes to the policy debates, by recourse to three issues that have been put on the agenda by growth theorists. One of the issues is that *human capital* formation may explain long term patterns of economic growth very well. (See, for instance, Lucas (1988), Tamura (1991), Glomm and Ravikumar (1992) or Caballé and Santos (1993).)

A second issue concerns *population*. For instance, Kremer (1993), Deardorff (1994) or Romer (1996) show that population size and its growth explain patterns of economic growth, if one looks at very long time horizons (Kremer's is more than one million years!). In those models the larger the population is, the more likely technological progress and so higher growth is. Furthermore, Becker, Murphy and Tamura (1990) or Rosenzweig (1990) show that economically driven fertility choices and human capital investments may lead to growth.

Thirdly, the paper considers the theory of *distribution* and *growth* which has been analyzed in a vast number of contributions. Just to name a recent few suffice it to mention Bertola (1993), Alesina and Rodrik (1994) or Garcia-Peñalosa (1995b) who derive interesting conclusions on the relationship between (re-)distribution and growth, mainly showing that the relationship is negative if resources are distributed to the non-accumulated factor of production.

The paper takes as its starting point the stylized fact that high growth economies have larger or more efficient *public education* systems and may show low degrees of income inequality. Even in countries such

[1] For discussions of this issue see e.g. Deininger and Squire (1996), Bénabou (1996), Bertola (1998) or Aghion and Howitt (1998), chpt. 9.

as the US a very significant fraction of education is carried out publicly.[2] The population in the model is made up of *high skilled* or *low skilled* workers. So the model argues that the population *composition* matters in the growth process by assuming that human capital can be identified with 'degrees'. Thus, education is taken to be '*lumpy*' in that people are hired as high skilled workers in the labour market only if they have obtained a degree.

In the paper the agents own the initial capital stock *equally.* In the literature on human capital investment, it is usually shown that given some distribution of innate abilities and costly education, optimizing agents sort themselves into high and low skilled workers depending on the path of the wage rates and the distribution of wealth. (See, for instance, Mincer (1958) or Findlay and Kierzkowski (1983).) If markets work perfectly, the sorting will see to it that the lifetime utility of a high and a low skilled worker is equalized, and that there is no inequality in the present value of lifetime earnings. However, should the capital market not function perfectly so that agents wishing to fund education cannot borrow against future earnings, the sorting will be distorted and one would observe inequality in the present value of lifetime earnings. The assumption of equal capital ownership eliminates that effect on the people's choice of education. Instead, in the paper the source of wage inequality lies in the production process itself. High skilled people carry human capital that enables them to perform all the tasks a low skilled person can do and *more. Effective labour* depends on *basic skills* and *high skills* in production. By assumption *basic skills and high skills are imperfect substitutes* in production, but *low and high skilled people are perfect substitutes in basic skills.* Thus, high skilled people may always perform the tasks of low skilled people, but low skilled people can never execute tasks that require a degree. In a perfectly competitive labour market this entails that the high skilled workers get a wage *premium* over and above what their low skilled colleagues receive.

[2] Notice that governments have fiscal and institutional instruments other than direct provision of education at their disposal that have a significant bearing on the working of private education systems. For a discussion of public vs. private education see Glomm and Ravikumar (1992).

The wage premium is shown to depend negatively on the percentage of high skilled people, which captures an important and realistic aspect in the explanation of wage rate inequality. (For empirical studies on this issue see, for instance, Freeman (1977), Katz and Revenga (1989), Bound and Johnson (1989), Tilak (1989) or Londoño (1990).)

The government's task is to provide and finance education. Respecting the right of private property, the government finances education by raising a tax on the wealth of all individuals.[3] So even those who have not received education contribute to financing it. That is realistic in most public education systems and - as shown below - is in the low skilled people's interest. The model postulates a simple relationship between government revenue and education output, by which the percentage of high skilled people in population is directly related to the tax rate. *Ex ante* every agent would like to and may get a degree so that innate ability differences are not important in the set-up.[4] That is so because I assume that education is provided as a public good.[5]

---

[3]The command optimum would involve expropriation of the capital stock. As that is not common in the real world, it is ruled out. For a similar point cf. Alesina and Rodrik (1994).

[4]In the model agents are endowed by some *basic* ability and receive basic education which is produced and provided costlessly. In the paper education is always meant to be higher education. *Ex ante* everybody is a candidate for receiving (higher) education and once chosen to be *in* the education process will complete the degree. The education process is taken to be sufficiently productive in converting no skills into high skills. Even if people have the same innate abilities and the same initial endowments and although the capital market functions perfectly, there is inequality in the present value of lifetime earnings in the model. For a recent model that studies the positive (and causal) link from income equality to human capital accumulation and high growth, see Chiu (1998). His model attributes the source of inequality to innate ability differences and liquidity constraints. This paper offers a different, technology based explanation with a positive (causal) link from human capital to income equality and high growth for given policy.

[5]I abstract from problems arising from the time spent receiving education. Of course, students forgo wage earnings for some time with the expectation that they are compensated by higher wages in the future. If one follows the human capital literature (e.g. Mincer (1958)) and views education as an effort demanding process, causing students to experience disutility while learning, the model may easily account for that by endowing the low skilled by some fixed positive and endowing the high skilled by some fixed negative amount of lifetime 'happiness', without altering the qualitative results of the paper. Along these lines I consider only adults who are at

In the market equilibrium growth is positively related to the percentage of high skilled people in population up to a certain point. That is so because the government takes resources away from the private sector in order to finance education, which reduces growth. On the other hand it generates more high skilled people which exert a positive effect on production, growth and wage equality. For maximum growth taxes and so the number of high skilled people must not be too high. Furthermore, growth and wage equality depend positively on the productivity of the education sector.

Next, a public policy analysis is conducted. If the government's welfare function attaches fixed weights to the representative high or low skilled individual's utility, a policy that only represents the *low skilled* worker is like a *Rawlsian* policy. Both choose the growth maximizing number of high skilled people in the model. The intuition for this is that the low skilled worker's wage does not depend on x, the percentage of high skilled people in population. But their capital income does depend on x, which explains why the low skilled worker chooses the growth maximizing x and the highest after-tax return on capital. A striking implication of the model is that growth maximization and *Rawlsian* preferences yield identical policies.

The opposite holds for a government that only represents the average *high skilled* worker. It acts like an *Anti-Rawlsian* government. Both choose x *lower* than the growth maximizing one, because the wage premium depends negatively on x. High skilled workers do not like too many of their own kind, since that reduces the wage premium but raises their capital income. In the optimum they choose a x lower than the growth maximizing one.

In the model a *strictly utilitarian* government faces the *non-trivial* problem of maximizing individual utility levels *and* the weight, it attaches to them. It optimally sets a x higher than the growth maximizing one,

least as old as the ones with a degree. Then the low skilled joining the labour pool as adults may be taken to have received utility by being idle during adolescence. In contrast, the high skilled would have studied and suffered disutility up to adulthood. Alternatively, one may simply assume that all people spent the same time in school, but attend different courses leading to different degrees.

which implies that it attaches more weight to having high skilled people in the economy than making the average high or low skilled individual 'happy'.

A comparison of the different policies indicates that a government serving the average low skilled worker chooses a growth maximizing policy. It is ambiguous whether in comparison to each other the utilitarian or the average high skilled labour serving government has higher growth, but both choose less than maximal growth.

The optimal low skilled workers' policy implies a more equitable wage distribution than a high skilled workers' one. In comparison, the strictly utilitarian government chooses a more equal wage ratio than the low skilled workers' government. Interestingly, a *strictly utilitarian* policy is *more egalitarian* in the model than a *Rawlsian* one.

High and low skilled labour form the clientele of representative governments. The public policy analysis suggests that the recent high growth, relatively low income inequality experience of some countries may be due to a combination of public education policies and improvements in the productivity of the education sector which is difficult to measure. If the productivity is equal across some countries, a policy favouring low skilled labour maximizes growth, attracts capital by granting high after-tax returns and reduces inequality compared to the optimal policy of high skilled labour.

The paper is organized as follows: Section 2 describes the economy. Section 2.1 derives the optimal behaviour of the private sector and the market equilibrium. Section 2.4 investigates optimal policies for governments with different welfare functions. Section 2.5 compares the optimal polices and section 3 provides some concluding remarks.

2 The Model

Consider an economy that is populated by N (large) members of two representative dynasties of infinitely lived individuals. The two dynasties are *high skilled* workers, L_1, and *low skilled* workers, L_0, where L_1, L_0

denote the total numbers of the respective agents in each dynasty. The difference between high and low skilled labour is "lumpy", that is, *either* an individual has received education in the form of a degree and is then considered high skilled *or* it has no degree and remains in the low skilled labour pool. By assumption the population is stationary so that $L_1 \equiv xN$ and $L_0 \equiv (1 \Leftrightarrow x)N$ where x denotes the percentage of high skilled people in population. The members of the dynasties supply one unit of either high or low skilled labour inelastically over time. Each worker initially owns an equal share of the total capital stock, which is held in the form of shares of many identical firms operating in a world of perfect competition. Thus, all agents receive wage and capital income and make investment decisions.[6] I assume that aggregate production is given by

$$Y_t = A_t \; K_t^{1-\alpha} \; H^\alpha, \;\; H^\alpha = \left[(L_1 + L_0)^\alpha + L_1^\alpha \right], \qquad 0 < \alpha < 1, \qquad (1)$$

where K_t denotes the aggregate capital stock including disembodied technological knowledge,[7] H measures effective labour in production, and A_t is a productivity index at time t. The production function is a reduced form of the following relationship between high and low skilled labour:[8] By assumption *effective labour* depends on *basic skills* and *high skills* and that *basic skills and high skills are imperfect substitutes* in production. On the other hand it is assumed that *low and high skilled people are perfect substitutes in basic skills.* Thus, high skilled people may always perform the tasks of low skilled people in the model, but low skilled people can never execute tasks that require a degree. Notice that each type

[6]Postulating an equal initial wealth (capital) distribution is a simplification, which serves to bring out clearly the effects of different policies on the income distribution. Alternatively, suppose a third type owns *all* the initial capital stock. Then one may verify that all the results for governments representing high or low skilled workers hold. Also, as will be shown below, the workers' utility depends on the balanced growth rate so that analyzing high and low skilled workers embodies the problem, a capital owning class would have.

[7]Thus, technological knowledge is taken to be a sort of capital good which is used to produce final output in combination with other factors of production. For an up-to-date discussion of these kinds of endogenous growth models see, for instance, Aghion and Howitt (1998), chpt. 1.

[8]For a more detailed explanation see Appendix A.

of labour alone is *not* an essential input in production.

The government runs a balanced budget at each point in time, uses its tax revenues to finance public education and maintains a constant ratio of expenditure G_t to its tax base.[9] It taxes the agents' wealth holdings at a constant rate τ on the capital stock of the representative agent $k_t = \frac{K_t}{N}$. So $G_t = \tau k_t N = \tau K_t$ and $\frac{G_t}{K_t} = \tau$ for all t. Thus, real resources are taken from the private sector and used to finance public education, which generates high skilled workers. In general, public education is 'produced' using government resources and other factors such as high skilled labour itself. That is captured by the following *reduced form* of the education technology

$$x = \tau^{\epsilon} \quad \text{where} \quad 0 < \epsilon \leq 1 \ , \tag{2}$$

$x_{\tau} = \epsilon\tau^{\epsilon-1} > 0$ and $x_{\tau\tau} = \epsilon(\epsilon \Leftrightarrow 1)\tau^{\epsilon-2} \leq 0$. Thus, if the government channels more resources into the education process, it will generate more high skilled people, $x_{\tau} > 0$. However, doing this generally becomes more difficult at the margin. This is supposed to reflect that, if $x_{\tau\tau} < 0$, more public resources provided to the education sector lead to a decreasing marginal product of those resources due to congestion or other effects. The parameter ϵ measures the productivity of the education sector.[10] If $\epsilon < 1$, the education sector is productive and a marginal increase in the wealth tax rate increases education output substantially. Underlying that is the description of an education sector with spillovers from, for instance,

[9] Various tax bases would be possible to investigate in the model. Capital taxes are considered to keep the analysis simple and are supposed to capture a broad class of tax arrangements. For a similar approach in a different context see Alesina and Rodrik (1994).

[10] The reduced form education technology directly relates the percentage of high skilled people (x) to the percentage of resources (wealth) going into the education sector (τ). Then ϵ may be viewed as the elasticity of education output to education financing (input) and may be interpreted as a productivity measure. As $x_{\epsilon} = \tau^{\epsilon}\ln(\tau) < 0$ because $\tau < 1$, a higher ϵ reduces output for given τ. Thus, for given policy a decrease in ϵ reflects a more productive education technology. Also and more conventionally, let $pr = \frac{x}{\tau}$ denote productivity. Then $pr = \tau^{\epsilon-1}$, which is decreasing in ϵ as well. Hence, for given policy productivity would be decreasing in ϵ according to both productivity concepts.

high skilled to new high skilled people or where the capital equipment such as computers makes the education technology very productive. The case $\epsilon = 1$ looks as though the education technology were quite productive as well, since then the number of high skilled people rises one-to-one with an increase in the tax rate. However, $x_\epsilon < 0$ for given policy so that a higher ϵ leads to less high skilled people (education output). Combining this with the assumption of a non-increasing marginal product ($x_{\tau\tau} \leq 0$) may justify calling $\epsilon = 1$ a relatively unproductive education technology.

Equation (2) is compatible with many models that also use high skilled labour as an input generating education. For instance, let h_t denote the *total* stock of human capital in the economy in a discrete time model. Following Azariadis and Drazen (1990) assume that human capital evolves according to

$$h_{t+1} = f(G_t, K_t, h_t)\, h_t$$

where new human capital h_{t+1} is produced by non-increasing returns. Here human capital formation would depend on the level of the stock of knowledge h_t, government resources provided for education G_t and the tax base K_t. The function $f(\cdot)$ governs the evolution of human capital. Assume that it is separable in the form $f(g(G_t, K_t), h_t)$. Let $g = c\left(\frac{G_t}{K_t}\right) = c(\tau)$ and for simplicity

$$h_{t+1} = c(\tau)\ h_t^\beta, \quad \text{where} \quad c \geq 0,\ c' > 0,\ c'' \leq 0,\ 0 < \beta < 1.$$

where β measures the productivity of the education sector and $c(\tau)$ captures the efficiency or quality of education, depending on the government resources channeled into education.[11] What distinguishes this model from those contributions is that in this paper human capital is carried discretely by the agents and so $h_t = x_t N$. Normalize population by setting $N = 1$. Then total human capital at date t is given by x_t. In a

[11] For a similar expression in an optimizing agent framework, see Nerlove, Razin, Sadka and von Weizsäcker (1993) eqn. (7).) That is a widely used specification. See, for example, eqns. (1), (2) in Glomm and Ravikumar (1992), eqn. (1) in Eckstein and Zilcha (1994), or eqn. (2) in Razin and Yuen (1996).

steady state $\bar{x} = x_t = x_{t+1}$ and so

$$\bar{x} = c(\tau)^{\frac{1}{1-\beta}}.$$

Next suppose that the efficiency of the education sector is described by $c(\tau) = \tau^{\mu}$ where $0 < \mu < 0$. For non-increasing returns to scale it is necessary that $\mu + \beta \leq 1$. Let $\frac{\mu}{1-\beta} \equiv \epsilon$ then the more explicit set-up would be equivalent to (2) in steady state. As $\bar{x}_{\epsilon} < 0$, any increase ϵ would mean that less human capital is generated in steady state. From the assumption of non-increasing returns to scale it follows that $\mu \leq 1 \Leftrightarrow \beta$ so that $\epsilon \leq 1$. Hence, $\epsilon = 1$ would represent a relatively unproductive human capital formation process.

Finally, notice that from equation (2) choosing τ is equivalent to choosing x. For the rest of the chapter it is convenient to use the inverse relationship $\tau = x^{\frac{1}{\epsilon}}$ whenever the government chooses taxes.

2.1 The Private Sector

There are as many identical firms as individuals and the firms face perfect competition. I assume that there is a capital spillover, which takes the form $A_t = \left(\frac{K_t}{N}\right)^{\eta} = k_t^{\eta}$, where $\eta \geq \alpha$, so that the *average* capital stock is the source of a positive externality.[12] Then simplify by setting $\eta = \alpha$ which allows one to concentrate on steady state behaviour. For a justification see Romer (1986). As the firms cannot influence the externality, it does not enter their decision directly so that

$$\begin{aligned} r &= (1 \Leftrightarrow \alpha) k_t^{\alpha} K_t^{-\alpha} H^{\alpha}, \\ w_1 &= \alpha k_t^{\alpha} K_t^{1-\alpha} \left[(L_1 + L_0)^{\alpha-1} + L_1^{\alpha-1} \right], \\ w_0 &= \alpha k_t^{\alpha} K_t^{1-\alpha} (L_1 + L_0)^{\alpha-1}. \end{aligned} \tag{3}$$

The workers have logarithmic utility and own all the assets which are collateralized one-to-one by capital. A representative worker takes the

[12] The model also works if the externality depends on the *entire* capital stock instead.

paths of r, w_1, w_0, τ as given and solves the problem

$$\max_{c_i} \int_0^{\infty} \ln c_i \; e^{-\rho t} \; dt \tag{4}$$

$$s.t. \qquad \dot{k} = w_i + (r \Leftrightarrow \tau)k \Leftrightarrow c_i \qquad i = 0, 1 \tag{5}$$

$$k_0 = \text{constant}, \;\; k_\infty = \text{free}.$$

Equation (5) is the worker's dynamic budget constraint. The worker's problem is a standard one (see e.g Chiang (1992), chpt. 9.) and its solution involves the following growth rate of the average high or low skilled worker's consumption

$$\gamma = \gamma_{c_0} = \gamma_{c_1} = (r \Leftrightarrow \tau) \Leftrightarrow \rho. \tag{6}$$

So consumption of all workers grows at the same rate in the optimum and depends on the after-tax return on capital. As the workers own the initial capital stock equally and have identical utility functions, their investment decisions are the same. Thus, the wealth distribution will not change over time and all agents continue to own equal shares of the total capital stock over time. The only difference in utility stems from different wage incomes which affect the instantaneous levels of steady state consumption.

2.2 Market Equilibrium

For the rest of the paper normalize by setting $N = 1$ so that the factor rewards in (3) are given by

$$r = (1 \Leftrightarrow \alpha)(1 + x^{\alpha}) \; , \qquad w_1 = \alpha k_t (1 + x^{\alpha - 1}) \qquad \text{and} \qquad w_0 = \alpha k_t. \tag{7}$$

The return on capital is constant over time and wages grow with capital. As $w_1 = w_0\,(1 + x^{\alpha-1})$, high skilled labour receives a premium over what their low skilled counterpart gets. That reflects the fact that the high skilled may always perfectly substitute for low skilled labour so that both types of labour receive the same wage w_0 for routine tasks and that performing high skilled tasks is remunerated by the additional amount

$w_0\, x^{\alpha-1}$. The premium depends on the percentage of high skilled labour in the population, grows over time at the rate γ and is decreasing in x for a given capital stock.

From the production function one immediately gets $\gamma_y = \gamma_k$ so that per capita output and the capital-labour ratio grow at the same rate. With constant N total output also grows at the same rate as the aggregate capital stock. From (6) the consumption of the representative worker grows at γ. Each worker owns $k_0 = \frac{K_0}{N}$ units of the initial capital stock. Equation (5) implies $\dot{k} = w_i + (r \Leftrightarrow \tau)k \Leftrightarrow c_i$ so that $\gamma_k = \frac{w_i - c_i}{k} \Leftrightarrow (r \Leftrightarrow \tau)$ for $i = 0, 1$ where $(r \Leftrightarrow \tau)$ is constant. In steady state, γ_k is constant by definition. But $\frac{w_i}{k}$ is constant as well, because from (7)

$$\frac{w_1}{k_t} = \frac{\alpha k_t (1 + x^{\alpha-1})}{k_t} = \alpha(1 + x^{\alpha-1}) \quad \text{and} \quad \frac{w_0}{k_t} = \alpha,$$

which implies $\gamma_k = \gamma$. Thus, the economy is characterized by *balanced growth* in *steady state* with $\gamma_Y = \gamma_K = \gamma_y = \gamma_k = \gamma_{c_1} = \gamma_{c_0}$.

Furthermore, from equation (5) and using $\gamma_k k = \dot{k}$ and $\gamma_k = \gamma_{c_1} = \gamma_{c_0}$ in steady state one obtains $(r \Leftrightarrow \tau \Leftrightarrow \rho)k_t = w_i + (r \Leftrightarrow \tau)k_t \Leftrightarrow c_i$. Thus,

$$c_1 = w_1 + \rho k_t \quad \text{and} \quad c_0 = w_0 + \rho k_t \tag{8}$$

are the instantaneous consumption levels of a representative high or low skilled worker in steady state.

From (6), (7) and $\tau = x^{\frac{1}{\epsilon}}$ one obtains $\gamma = (1 \Leftrightarrow \alpha)\,(1 + x^{\alpha}) \Leftrightarrow x^{\frac{1}{\epsilon}} \Leftrightarrow \rho$ so that for given τ an increase in x raises growth. The necessary first order condition for growth maximization involves $\alpha(1 \Leftrightarrow \alpha)x^{\alpha-1} = \frac{x^{\frac{1}{\epsilon}-1}}{\epsilon}$ which upon solving for x establishes that

$$\hat{x} = [\epsilon\alpha(1 \Leftrightarrow \alpha)]^{\frac{\epsilon}{1-\epsilon\alpha}} \tag{9}$$

is the growth maximizing percentage of high skilled workers in the population and $\hat{\tau} = [\epsilon\alpha(1 \Leftrightarrow \alpha)]^{\frac{1}{1-\epsilon\alpha}}$ is the growth maximizing tax rate.

Lemma 1 *A growth maximizing government chooses* $\hat{x} = [\epsilon\alpha(1 \Leftrightarrow \alpha)]^{\frac{\epsilon}{1-\epsilon\alpha}}$

and $\hat{\tau} = [\epsilon\alpha(1 \Leftrightarrow \alpha)]^{\frac{1}{1-\epsilon\alpha}}$.

Growth is a concave function of x since for $\epsilon \leq 1$ and any x

$$\frac{d^2\gamma}{(dx)^2} = \Leftrightarrow \alpha(1 \Leftrightarrow \alpha)^2 x^{\alpha-2} \Leftrightarrow \frac{1}{\epsilon}\left(\frac{1}{\epsilon} \Leftrightarrow 1\right) x^{\frac{1-2\epsilon}{\epsilon}} < 0$$

and the marginal growth rate for $x \to 0$ is infinity,

$$\lim_{x \to 0} \frac{d\gamma}{dx} = x^{\alpha-1}\left[\alpha(1 \Leftrightarrow \alpha) \Leftrightarrow \frac{x^{\frac{1}{\epsilon}-\alpha}}{\epsilon}\right] = +\infty$$

since $\frac{1}{\epsilon} \geq 1$ by assumption. By the concavity of γ and given the above properties there is a $\gamma(x)$, generating the same growth as $\gamma(0)$. In that case the government chooses a rather high tax rate, implying a high x. So in the model it is possible that an economy has high skilled workers, but does not do better than another economy with no high skilled people. That x is given by

$$\begin{aligned} \gamma(0) \Leftrightarrow \gamma(x) &= (1 \Leftrightarrow \alpha) \Leftrightarrow \rho \Leftrightarrow (1 \Leftrightarrow \alpha)\left[1 + x^{\alpha}\right] + x^{\frac{1}{\epsilon}} + \rho = 0 \\ \tilde{x} &= (1 \Leftrightarrow \alpha)^{\frac{\epsilon}{1-\epsilon\alpha}} \end{aligned} \tag{10}$$

and clearly $\tilde{x} > \hat{x}$. The effect of a change in the productivity of the education sector for a given $x \in (0,1)$ is given by

$$\frac{d\gamma}{d\epsilon} = \frac{\ln(x)\ x^{\frac{1}{\epsilon}}}{\epsilon^2} < 0.$$

An increase in ϵ makes the production of education *more difficult*. So a reduction in ϵ, that is, making the education technology more productive, raises the growth rate. Hence, the growth maximizing x must also be higher.

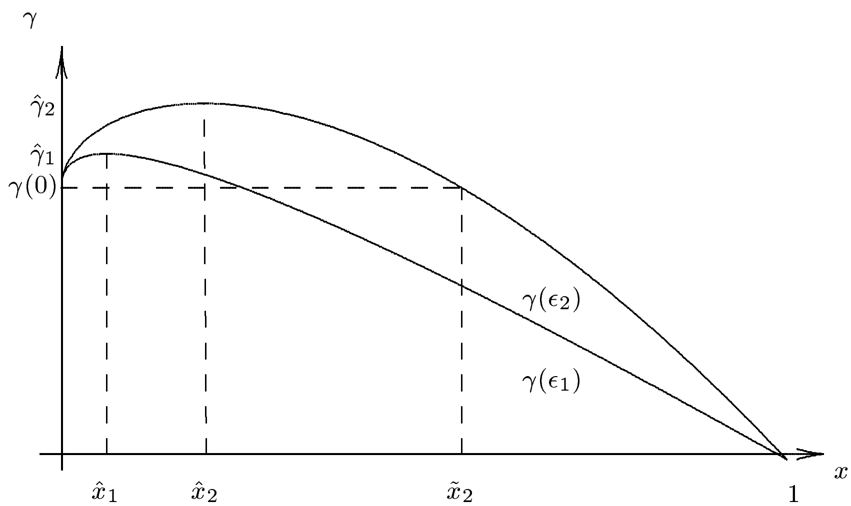
Lemma 2 *The growth rate γ has the following properties:*

1. γ is concave in x. *2.* $\lim_{x \to 0} \frac{d\gamma}{dx} = +\infty$.

3. $\frac{d\gamma}{d\epsilon} < 0$ *for* $x \in (0,1)$. *4. If* $\tilde{x} = (1 \Leftrightarrow \alpha)^{\frac{\epsilon}{1-\epsilon\alpha}}$, *then* $\gamma(0) = \gamma(x)$.

The properties can be read off from Figure 1. It can be seen that the

Figure 1: γ as a function of x for different ϵ



Parameter values: $\alpha = \frac{1}{2}$, $\rho = 0.01$, $\epsilon_1 = 1$, $\epsilon_2 = 0.6$.

growth maximizing x increases with an increase in the productivity (lower ϵ) of the education technology and that there exists a $\tilde{x}$ where $\gamma(0) = \gamma(x)$.

2.3 Income Inequality

In the model all income differences between individuals are due to differences in wage income. If one wants to relate growth to income inequality it makes sense to look at an average of personal wage incomes over time. If the agents can sell their income stream in a perfect capital market, they will discount their stream of wages by the market rate of return on assets, $r \Leftrightarrow \tau$. Consequently, the present value of their lifetime wages is

$$\int_0^\infty w_{it}\, e^{-(r-\tau)t} dt = \int_0^\infty w_{i0}\, e^{\gamma t}\, e^{-(r-\tau)t} = \frac{w_{i0}}{\rho} \equiv w_i^d \quad \text{where} \quad i = 0, 1.$$

Thus, w_i^d denotes the sum of an individual's wage income discounted by the market rate of return on assets. That is the income concept used when analyzing the wage income distribution in this paper.[13] Notice

$$w_0^d = \frac{w_0}{\rho} = \frac{\alpha k_0}{\rho} \quad \text{and} \quad w_1^d = \frac{w_1}{\rho} = \frac{\alpha k_0(1 + x^{\alpha-1})}{\rho} \tag{11}$$

and that the mean of the discounted sum of wage incomes is

$$\mu^d = (1 \Leftrightarrow x)w_0^d + xw_1^d = \frac{(1 + x^{\alpha})\alpha k_0}{\rho}. \tag{12}$$

implying $\frac{dw_0^d}{dx} = 0$, $\frac{dw_1^d}{dx} < 0$ and $\frac{d\mu^d}{dx} > 0$, that is, the mean of the PV of lifetime wage income is increasing in x. In order to compare any two cumulative distribution functions of discounted lifetime wage income assume $x_1 > x$. Then the different values of x will give rise to two cumulative distribution functions, $F(w_i^d(x_1))$ and $G(w_i^d(x))$, which have unequal means.

If F dominates G in the sense of *Second Order Stochastic Dominance* (SOSD), then F will be preferred to G by any increasing, concave social welfare function according to Atkinson (1970). Geometrically, a distribution $F(w)$ dominates another distribution $G(w)$ in the sense of SOSD if over every interval $[0, c]$, the area under $F(w)$ is never greater (and sometimes smaller) than the corresponding area under $G(w)$.[14]

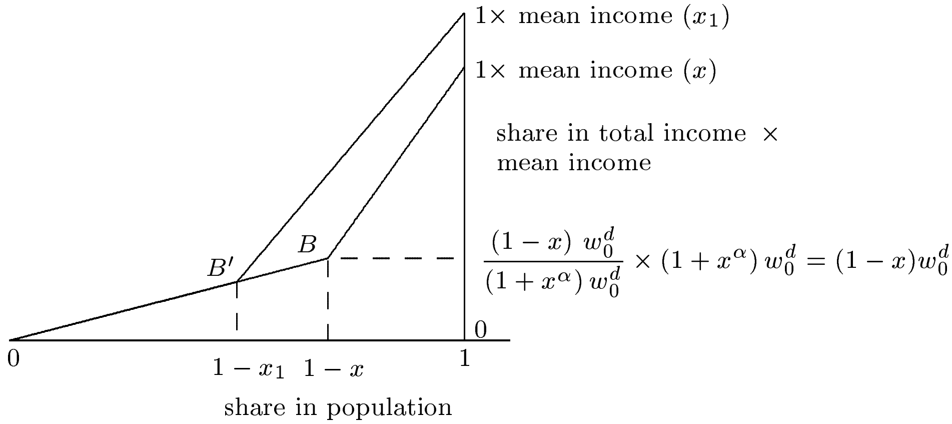
Second Order Stochastic Dominance is equivalent to *Generalized Lorenz Curve* (GLC) dominance. (For a proof see, for example, Lambert (1993), pp. 62-66.) A GLC is obtained by multiplying the values of the y-axis of an ordinary Lorenz Curve, which relates the share of the population (x-axis) to the share in total income (y-axis) that that population share receives, by mean income, i.e. (share of total income) × (mean income). The GLC for the distribution of the PV of lifetime wages is

[13] Other income variables one may want to use are current wage income w_{it}, detrended initial wages w_{i0}, or capital adjusted wages $\frac{w_{it}}{k_t}$. All of these concepts suffer from the problem that do not fully reflect the path the wages follow.

[14] The concept derives from evaluating risky returns under conditions of uncertainty. See e.g. Hirshleifer and Riley (1992), chpt. 3.4. Formally and for non-negative incomes, *Second Order Stochastic Dominance* requires $\int_0^c F(w)dw \leq \int_0^c G(w)dw$.

presented below.

Figure 2: Generalized Lorenz Curve



A GLC *dominates* another one if the two curves do not cross and one is completely above the other one. In the figure the income distribution associated with $x_1 > x$ GLC-dominates the income distribution for x. The reason is that an increase in x raises μ^d and shifts the kink at B to a point B' which is to the left and on the old GLC(x).

According to a theorem by Shorrocks (1983) GLC dominance carries with it welfare approval according to every increasing, strictly concave utility-of-income function. Put another way: Every individualistic additively separable symmetric and inequality-averse social welfare function would prefer the GLC dominating income distribution. That means that according to the GLC dominance criterion there exists a *unanimous* preference for the (PV of lifetime wage) income distribution with the higher GLC. Even the high skilled would prefer the distribution with a higher x under e.g. a veil of ignorance. In that sense the model's society as a whole would prefer the distribution of the PV of lifetime wage income generated by the higher percentage of high skilled people.[15]

[15] It is straightforward to see that exactly the same holds for the distribution of detrended (initial) wage incomes w_{i0} and capital adjusted wages $\frac{w_{it}}{k_t}$. It also holds if one works with current wage rates w_{it} and $x \leq \hat{x}$. In that case an increase in x causes the new GLC to be everywhere above the old GLC for $t > 0$, because the capital stock would be higher at each date and mean income would rise. However, if $x > \hat{x}$ it does not necessarily hold.

Let $I(x)$ be any inequality measure reflecting that a higher x leads to a GLC dominating distribution.[16] Then $I(0) = I(1) = 0 < I(x)$ and $\frac{dI}{dx} < 0$ for $x \in (0, 1)$. Thus, according to $I(x)$ and for the PV of lifetime wage incomes there is no measured inequality if all agents get the same wage and they are all either equally high or low skilled. If there is any skill heterogeneity, producing more skills reduces inequality in the PV of lifetime wage incomes if measured by $I(x)$. Furthermore, as $x = \tau^{\epsilon}$ and so $I(\tau)$, a decrease in ϵ for a given policy τ, would lower $I(x)$.

Proposition 1 *If there is heterogeneity in skills, $x \in (0, 1)$, an increase in the percentage of high skilled people or an increase in the productivity of the education technology (lower ϵ) for given policy reduce inequality in the present value of lifetime wage incomes in the sense of Generalized Lorenz Curve Dominance.*

Thus, according to the proposition and in terms of the PV of lifetime wage income an increase in the number of high skilled people represents an equalizing income transfer from a rich, high skilled to a relatively poor, low skilled person.

2.4 The Government

The government takes the optimal decision of the workers as given and chooses taxes to generate education output. This is equivalent to choosing x in the model. I assume that the government has different objectives. For instance, it may represent only high skilled workers, only low skilled workers or a mixture of the two. Integrating the utility of the representative low and high skilled worker as given by (4) one obtains (See

[16] A simple measure satisfying the properties of $I(x)$ is $I = \frac{w_1^d}{w_0^d} - 1$ which Fields (1987), axiom A5, calls *relative gap inequality*, defined in terms of mean wage incomes of the two groups. Notice that the measure as such does not take explicit account of the population composition. However, in the model it implicitly does, because the wage rates depend on the relative number of high skilled people. In appendix B I check the properties of $I(x)$ against those of some other commonly used inequality measures. It is shown that, for instance, the variance and the coefficient of variation have the properties of $I(x)$ if $\alpha \leq \frac{1}{2}$.

Appendix C.)

$$V^h = \frac{\ln c_1}{\rho} + \frac{\gamma}{\rho^2} = \frac{\ln\left((\alpha(1+x^{\alpha-1})+\rho)k_0\right)}{\rho} + \frac{\gamma}{\rho^2} \quad \text{and} \tag{13}$$

$$V^l = \frac{\ln c_0}{\rho} + \frac{\gamma}{\rho^2} = \frac{\ln\left((\alpha+\rho)k_0\right)}{\rho} + \frac{\gamma}{\rho^2}. \tag{14}$$

Superscript h (l) stands for high (low) skilled. Notice that $V^h > V^l$ for $x \in (0,1)$.

A Class of Governments. Consider a government's social welfare function $W^b(V^h, V^l) = \zeta\ V^h + (1 \Leftrightarrow \zeta)\ V^l$ with $\zeta \in [0,1]$ which attaches fixed weights on the individual high and low skilled agent's utility. If $\zeta = 1$ the government is only concerned about the welfare of the representative high skilled worker and if $\zeta = 0$ it cares about the average low skilled worker only. For all other values of ζ it represents a mixture of the representative agents' utility. The FOC for the maximization of W^b is given by

$$\zeta\frac{\partial V^h}{\partial x} + (1 \Leftrightarrow \zeta)\frac{\partial V^l}{\partial x} = \frac{\zeta}{\rho}\left(\frac{\partial c_1}{\partial x}\frac{1}{c_1} + \frac{1}{\rho}\frac{\partial \gamma}{\partial x}\right) + \frac{(1 \Leftrightarrow \zeta)}{\rho}\left(\frac{\partial c_0}{\partial x}\frac{1}{c_0} + \frac{1}{\rho}\frac{\partial \gamma}{\partial x}\right) = 0.$$

Notice that $\frac{\partial c_0}{\partial x} = 0$ because low skilled labour's wages and consumption do not depend on x. Simplification yields

$$\Leftrightarrow \left(\frac{\zeta}{\rho}\right)\left(\frac{\alpha(1 \Leftrightarrow \alpha)k_0 x^{\alpha-2}}{\alpha k_0(1+x^{\alpha-1}) + \rho k_0}\right) + \frac{\gamma_x}{\rho^2} = 0 \tag{15}$$

where $\gamma_x = \frac{\partial \gamma}{\partial x}$. From this one immediately obtains an important result. As the first expression on the LHS is negative for $\zeta > 0$, γ_x must be positive. Given the concavity of γ the government would choose $x < \hat{x}$. Thus, a government attaching positive weight to a representative high skilled worker chooses a smaller than the the growth maximizing x.

The case $\zeta = 0$ is of special interest because it is equivalent to the choice of a *Rawlsian* government. A Rawlsian government has a welfare function $W = min(V^h, V^l)$. As the wages of the high skilled are always greater than those of the low skilled, $c_1 > c_0$ and so $V^h > V^l$ for all $x \in (0,1)$. But then $\zeta = 0$ captures the preferences of a Rawlsian government with *leximin* preferences over the individuals' utilities. Thus, a Rawlsian government, which maximizes the utility of the *least well-off*, and a government representing the average low skilled worker, set the growth maximizing tax rate, $x_l = \hat{x}$.

If $\zeta = 1$ the government acts in the interest of the average high skilled worker. That government's choice is equivalent to an *Anti-Rawlsian* government with *leximax* preferences such that $W = max(V^h, V^l)$. Recall $\gamma_x = \alpha(1 \Leftrightarrow \alpha)x^{\alpha-1} \Leftrightarrow \frac{x^{\frac{1}{\epsilon}-1}}{\epsilon}$ and use (15) to get the FOC for $\zeta = 1$

$$\begin{aligned}
\frac{\epsilon\rho\alpha(1 \Leftrightarrow \alpha)k_0 x^{\alpha-2}}{\alpha k_0(1 + x^{\alpha-1}) + \rho k_0} &= \epsilon\alpha(1 \Leftrightarrow \alpha)x^{\alpha-1} \Leftrightarrow x^{\frac{1}{\epsilon}-1} \\
\frac{\epsilon\rho\alpha(1 \Leftrightarrow \alpha)}{\alpha(1 + x^{\alpha-1}) + \rho} &= \epsilon\alpha(1 \Leftrightarrow \alpha)x \Leftrightarrow x^{\frac{1}{\epsilon}-1+2-\alpha} \\
\left(\epsilon\alpha(1 \Leftrightarrow \alpha) \Leftrightarrow x^{\frac{1}{\epsilon}-\alpha}\right)\left(\alpha\left(x + x^{\alpha}\right) + \rho x\right) &= \rho\epsilon\alpha(1 \Leftrightarrow \alpha).
\end{aligned}$$

Notice $\hat{x}^{\frac{1}{\epsilon}-\alpha} = \epsilon\alpha(1 \Leftrightarrow \alpha)$ so that

$$\left(\hat{x}^{\frac{1}{\epsilon}-\alpha} \Leftrightarrow x^{\frac{1}{\epsilon}-\alpha}\right)\left(\alpha\left(x + x^{\alpha}\right) + \rho x\right) = \rho\hat{x}^{\frac{1}{\epsilon}-\alpha}. \qquad (16)$$

Hence, an increase in ζ makes a government choose a x lower than $\hat{x}$. The lowest x will be chosen by a government representing the representative high skilled worker. Call the x chosen by a $\zeta = 1$ government x_h. Then $\hat{x} > x > x_h$ for $\zeta \in (0,1)$.

Proposition 2 *If $\zeta = 0$, the government represents the average low skilled worker only and acts like a* <u>*Rawlsian*</u> *government. Both choose $\hat{x}$ and maximize the after-tax return on capital and growth.*

If $\zeta = 1$, the government represents the average high skilled worker only and acts like an <u>*Anti-Rawlsian*</u> *government. Both will set $x_h < \hat{x}$ and have lower growth than the* <u>*Rawlsian*</u> *government.*

Any other government with a welfare function W^b and $\zeta > 0$ will set $x < \hat{x}$. The optimal choices imply $\hat{\gamma}$ if $\zeta = 0$ and $\gamma(x_h)$ if $\zeta = 1$

where $\hat{\gamma} > \gamma(x_h)$. *Also for* $\zeta \in (0,1)$, $\gamma(x_h, 1) < \gamma(x, \zeta) < \gamma(\hat{x}, 0)$ *and* $I(1) > I(\zeta) > I(0)$.

Thus, wage inequality is lower and growth higher under a Rawlsian than under any other government with a welfare function W^b. Notice that the proposition does not claim that the Rawlsian government chooses to eliminate all inequality.[17]

A Strictly Utilitarian Government The strictly utilitarian government maximizes $W^u(V^h, V^l) = x\ V^h + (1 \Leftrightarrow x)\ V^l$, and its problem is *non-trivial* in the model, because maximization of W^u does not only involve maximizing the individual utility indices, but also choosing the weights $x, (1 \Leftrightarrow x)$ attached to them. Using (14) one may express $W^u(\cdot)$ as

$$\begin{aligned} W^u(x) &= \frac{x\ \ln c_1}{\rho} + \frac{x\ \gamma}{\rho^2} + \frac{(1 \Leftrightarrow x)\ \ln c_0}{\rho} + \frac{(1 \Leftrightarrow x)\ \gamma}{\rho^2} \\ &= \frac{x\ \ln c_1}{\rho} + \frac{(1 \Leftrightarrow x)\ \ln c_0}{\rho} + \frac{\gamma}{\rho^2}. \end{aligned}$$

The derivative of W^u with respect to x is given by

$$\frac{\partial W^u}{\partial x} \equiv \upsilon(x) = \frac{1}{\rho}\left(\ln c_1 + \frac{\partial c_1}{\partial x}\frac{x}{c_1} \Leftrightarrow \ln c_0 + \frac{\partial c_0}{\partial x}\frac{1 \Leftrightarrow x}{c_0}\right) + \frac{\gamma_x}{\rho^2}$$

where $\upsilon(x)$ denotes marginal welfare. As the initial consumption of the low skilled does not depend on x in steady state ($\frac{\partial c_0}{\partial x} = 0$) simplify to obtain

$$\upsilon(x) = \frac{1}{\rho}\left(\ln\left(\frac{c_1}{c_0}\right) + \frac{\partial c_1}{\partial x}\frac{x}{c_1}\right) + \frac{\gamma_x}{\rho^2}. \tag{17}$$

[17] Minimal inequality is not possible in the model unless $x = 0$ or $x = 1$ which is what a strictly utility or income egalitarian government chooses. These choices would also be Rawlsian. In Rehme (1998) it is shown that such an indeterminate policy is generally bad for growth. Interestingly, if the agents are very patient, they prefer $x = 0$. For the sake of realism this paper assumes from now on that there is always some heterogeneity in skills, $x \in (0,1)$.

For an optimum $\upsilon(x) = 0$ is required. Recall $c_1 = \alpha k_0(1 + x^{\alpha-1}) + \rho k_0$ so that

$$\Delta_1(x) \equiv \frac{\partial c_1}{\partial x}\frac{x}{c_1} = \Leftrightarrow \frac{\alpha(1 \Leftrightarrow \alpha)x^{\alpha-1}}{\alpha(1 + x^{\alpha-1}) + \rho} < 0.$$

Also for $\ln\left(\frac{c_1}{c_0}\right)$ one verifies that

$$\Delta_2(x) \equiv \ln\left(\frac{c_1}{c_0}\right) = \ln\left(\frac{\alpha(1 + x^{\alpha-1}) + \rho}{\alpha + \rho}\right) > 0$$

because $\frac{\alpha(1+x^{\alpha-1})+\rho}{\alpha+\rho} > 1$. Thus, unless $S(x) \equiv \Delta_2(x) + \Delta_1(x) = 0$ one gets $x \neq \hat{x}$ in (17). This follows from the concavity of γ. In appendix D I show that $S(x) > 0$ and strictly decreasing for all $x \in [0, 1]$. But then $\upsilon(x) > 0$ at $x = \hat{x}$ and so by the concavity of γ one must have $\gamma_x < 0$ in an optimum and $x_u > \hat{x}$ where subscript u denotes the optimal choice of the utilitarian government.

Furthermore, it is shown in appendix E that the level of welfare $W(x)$ is lower at $x = 1$ than at $\hat{x}$. But $W^u(\hat{x}) > W^u(1)$ implies $\upsilon(x = 1) < 0$ so that the optimal solution must satisfy $\hat{x} < x_u < 1$, because $S(x)$ is strictly decreasing for any $x > \hat{x}$ and γ is concave in x. As $W^u(\hat{x}) > W^u(1)$ and all the derivatives exist, there must be one $x \in (\hat{x}, 1)$ where $\upsilon(x) = 0$. Then, the choice $\hat{x} < x_u < 1$ implies $I(1) = 0 < I_u < I(\hat{x})$ and $\gamma(1) < \gamma_u < \hat{\gamma}$.

Proposition 3 *A strictly utilitarian government chooses* x_u *such that* $x_u \in (\hat{x}, 1)$ *implying* $0 < I_u < I(\hat{x})$ *and* $\gamma(1) < \gamma_u < \hat{\gamma}$.

This is an interesting result and the intuition for it is not as straightforward as it seems. The strictly utilitarian government maximizes the individual utility indices and the weights the groups contribute to overall welfare. More precisely, it trades off a higher individual high skilled worker's utility requiring a low x with its desire to maximize the number of high skilled people. On the other hand, it trades off higher welfare of each low skilled person which would imply choosing $\hat{x}$ with its desire to minimize the number of low skilled people. In the optimum, it attaches more welfare weight on having high skilled people in the economy than choosing the growth maximizing number of high skilled persons. Thus, it chooses lower than maximum growth, but by this it pushes down wage inequality compared to a growth maximizing government.

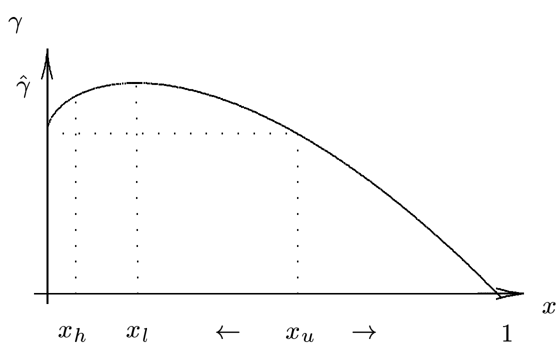
2.5 Comparison of the Different Policies

The propositions imply $0 < x_u, x_h, x_l < 1$ where x_u denotes the strictly utilitarian, x_h the high skilled labour and x_l the low skilled labour government's optimal choice. Thus,

Proposition 4 *The optimal policies of the governments are such that*

1. $\gamma_u, \gamma_h, \gamma_l > \gamma(1)$. *2.* $\gamma_l = \hat{\gamma} > \gamma_u, \gamma_h$. *3.* $\gamma_u \gtreqless \gamma_h$. *4.* $I_h > I_l > I_u$.

Figure 3: The governments' optimal policies



A government representing the average low skilled worker acts growth maximizing and generates less wage inequality than a government representing the average high skilled worker. Due to the externality, high skilled people exert on production, the low skilled workers' wages do not depend on x in equilibrium. Thus, they choose maximum growth and the highest after-tax return on capital.

For the average high skilled labour representing government things are quite different. The positive externality in production has a negative impact on their wages and that makes them choose a low x. On the other hand, a high x increases their capital income. The trade-off is solved so that the wage income component of their utility dominates and they do not choose the growth maximizing x. So it is socially desirable to have sufficient high skilled labour, but for the representative high skilled person it is privately bad, if too many of its kind are present.

The strictly utilitarian government maximizes the sum of the individual utility indices. It chooses *more* than the growth maximizing number of high skilled workers because it weighs the number of individuals more than the average utility of each type. That leads to a policy inducing a more equal wage income distribution. From the analysis it is not clear whether a utilitarian government has higher or lower growth than a government representing high skilled labour, but it will definitely have less income inequality.[18]

Interestingly, the strictly *utilitarian* government chooses a policy that is *more egalitarian* in terms of wages than a government representing low skilled labour and so more egalitarian than a *Rawlsian* policy. That provides an example that a Rawlsian objective does not always imply more egalitarianism than a utilitarian objective.[19]

3 Conclusion

The experience of high growth economies suggests that there is a positive link from providing education to income equality and growth. The paper offers a theoretical explanation of this stylized fact. Assuming that only people with a degree can take high skilled jobs, I show that the public choice of human capital directly affects income inequality and economic growth.

[18]This follows since the implicit solutions for the Anti-Rawlsian and the utilitarian governments in (16), resp. $v(x) = 0$ in (17) are not easily solved and depend in a non-linear way on the parameters of the model. As an exact solution does not add significantly to the qualitative results, I leave it an open question.

[19]The textbook comparison of utilitarian and leximin welfare functions usually argues that the choice of a utilitarian leads to *more* and not less inequality. See, for instance, Mas-Colell, Whinston and Green (1995), p. 828.

Due to market imperfections or institutional restrictions, the high skilled contribute more to effective labour in production than their unskilled counterpart. Hence, the number of people carrying high skills plays a crucial role in the model. The government raises taxes on all individuals and provides public education which produces human capital in the form of high skilled people. It is shown that the productivity of the education sector has a positive influence on growth and income equality. A (Rawlsian) government representing the average *unskilled* worker chooses the *growth maximizing* number of high skilled people. The average *high skilled* worker's government chooses less skilled people and makes the wages *more unequal* and growth *lower* than the Rawlsian policy. A strictly *utilitarian* government is shown to choose more high skilled people than the low skilled labour's government.

The paper's main insight lies in the result that a government representing the average *low skilled* worker chooses maximum growth, the highest after-tax return on capital and a wage distribution that is more equitable than the one chosen by a government representing the average *high skilled* worker. That stresses the importance of education policies in the growth process and their distributional consequences. It also provides a theoretical explanation why some highly competitive East Asian countries have relatively low income inequality and high growth.

Of course, it would be desirable to know more about the exact link between government revenues channelled into education and the education output. The level of human capital that individuals carry is clearly important. Human capital acquisition may entail more than one degree for different levels of human capital. These and other questions are left for further research.

Appendix

A Technology

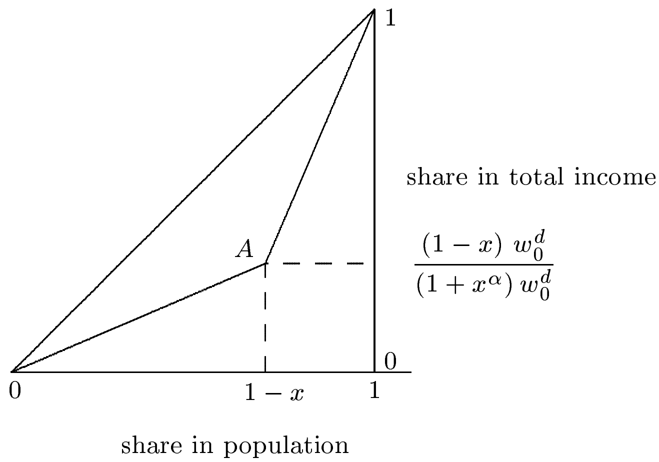
By assumption $Y_t = A_t H_t^{\alpha} K_t^{1-\alpha}$, where the index of effective labour H depends on labour requiring *basic skills* (B) and labour requiring *high skills* (S). *Labour* requiring basic skills is performed by high and low skilled *persons*, $B = B(L_0, L_1)$, whereas high skilled labour is only performed by high skilled persons, $S = S(L_1)$. High and low skilled people are perfect substitutes to each other when performing basic skill (routine) tasks, i.e. $B(L_0, L_1) = L_0 + L_1$. Thus, high skilled people also perform those routine tasks a low skilled person may do. On the other hand, only high skilled people can perform high skilled tasks (labour) and for simplicity let $S(L_1) = L_1$. To capture the relationship between labour inputs assume $H = [B^{\rho} + S^{\rho}]^{\frac{1}{\rho}} = [(L_1 + L_0)^{\rho} + L_1^{\rho}]^{\frac{1}{\rho}}$. For $\rho < 1$ labour requiring basic skills (B) and labour requiring high skills (S) are imperfect (less than perfect) substitutes. For ease of calculations let $\rho = \alpha < 1$ which yields equation (1). For a similar set-up in a different context see Garcia-Peñalosa (1995a).

B The relationship between $I(x)$ and some measures of income inequality

In this appendix and for convenience, the PV of lifetime wage income will simply be called (wage) income.

Lorenz Curve. A Lorenz Curves (LC) relates the population to the income shares. Total wage income is $\mu^d N$. Furthermore, $L_0 = xN$, $L_1 = xN$ and mean income μ^d is increasing in x. The share of total income going to the low skilled is $\frac{w_0^d L_0}{\mu^d N} = \frac{(1-x)w_0^d}{\mu^d}$ so that the Lorenz curve looks like Figure 4 below.

Figure 4: Ordinary Lorenz Curve



The Lorenz Curve (LC) has a kink at the point A at which $(1 \Leftrightarrow x)$ percent of the population receive $\frac{1-x}{1+x^\alpha}$ percent of total wage income. On the margin an increase in x shifts A to the left by 1 unit *for a given income share.* On the other hand, a marginal increase in x *reduces* the income share by

$$\frac{(1+x^\alpha)+\alpha x^{\alpha-1}(1 \Leftrightarrow x)}{(1+x^\alpha)^2}$$

for a given population share. If such a change would move A up to any new position above the old Lorenz Curve (LC dominance), then inequality would unambiguously have been reduced. Thus, one must analyze whether the movement of A to the left is greater or less than the movement up or down. In the model A moves down. Thus, the condition amounts to

$$(LHS\ \downarrow): \quad \frac{(1+x^\alpha)+\alpha x^{\alpha-1}(1 \Leftrightarrow x)}{(1+x^\alpha)^2} < 1 \quad : (RHS\ \leftarrow)$$

If x is rather low ($x \to 0$), the inequality does not hold. Hence, in general the LCs cross no unambiguous ranking of the wage income distributions is possible according to the LC dominance criterion.

Gini Coefficient. From the LC one may calculate the *Gini* coefficient as

$$G = 1 \Leftrightarrow 2\left[\frac{(1 \Leftrightarrow x)^2}{2(1+x^\alpha)} + \frac{x(1 \Leftrightarrow x)}{1+x^\alpha} + \frac{x^2(1+x^{\alpha-1})}{2(1+x^\alpha)}\right] = \frac{x^\alpha(1 \Leftrightarrow x)}{1+x^\alpha}$$

where the expression in square brackets represents the area under the LC. Then

$$\begin{aligned} sgn(G_x) &= \left[\alpha x^{\alpha-1}(1 \Leftrightarrow x) \Leftrightarrow x^\alpha\right](1+x^\alpha) \Leftrightarrow \alpha x^{\alpha-1} x^\alpha (1 \Leftrightarrow x) \\ &= x^{\alpha-1}\left([\alpha(1 \Leftrightarrow x) \Leftrightarrow x](1+x^\alpha) \Leftrightarrow \alpha x^\alpha(1 \Leftrightarrow x)\right) \end{aligned}$$

For low x, $x \to 0$, an increase in x raises the Gini index, whereas for higher values of x a higher x reduces it. Hence, the Gini coefficient does not produce unambiguous rankings of the wage income distribution.

The Variance and the Coefficient of Variation. The variance of personal wage income $V^d(x)$ is

$$\begin{aligned} V^d &= (w_1^d \Leftrightarrow \mu^d)^2\, x + (w_0^d \Leftrightarrow \mu^d)^2\,(1 \Leftrightarrow x) \\ &= \left[w_0^d(1+x^{\alpha-1}) \Leftrightarrow w_0^d(1+x^\alpha)\right]^2 x + \left[w_0^d \Leftrightarrow w_0^d(1+x^\alpha)\right]^2 (1 \Leftrightarrow x) \\ &= w_0^{d^2}\left[x^{2\alpha}\left(\frac{1 \Leftrightarrow x}{x}\right)^2 x + (\Leftrightarrow x^\alpha)^2\,(1 \Leftrightarrow x)\right] \\ &= w_0^{d^2}\, x^{2\alpha-1}(1 \Leftrightarrow x) \end{aligned}$$

which is decreasing in x if $\alpha < \frac{1}{2}$. The coefficient of variation is defined by $C^d \equiv \frac{\sqrt{V^d}}{\mu^d}$ and amounts to

$$C^d(x) = \frac{w_0^d\, x^\alpha \left(\frac{1}{x} \Leftrightarrow 1\right)^{\frac{1}{2}}}{w_0^d\,(1+x^\alpha)}.$$

The sign of $\frac{dC}{dx}$ depends on

$$\left[\alpha x^{\alpha-1}\left(\frac{1 \Leftrightarrow x}{x}\right)^{\frac{1}{2}} \Leftrightarrow \frac{x^\alpha}{2}\left(\frac{1 \Leftrightarrow x}{x}\right)^{-\frac{1}{2}} x^{-2}\right](1+x^\alpha) \Leftrightarrow \alpha x^{\alpha-1} x^\alpha \left(\frac{1 \Leftrightarrow x}{x}\right)^{\frac{1}{2}}$$
$$= \quad \alpha x^{\alpha-1}\left(\frac{1 \Leftrightarrow x}{x}\right)^{\frac{1}{2}}\left[\left(1 \Leftrightarrow \left(\frac{1 \Leftrightarrow x}{x}\right)^{-1}\left(\frac{1}{2\alpha x}\right)\right)(1+x^\alpha) \Leftrightarrow x^\alpha\right]$$

which is definitely negative for all $x \in [0,1]$ if $\alpha \leq \frac{1}{2}$.

C Welfare Measures

The workers' welfare integral is given by $U_t^j = \int_0^t \ln c_{j,t}\ e^{-\rho t} dt$ where $j = 0, 1$. Let $t \to \infty$ and use integration by parts. Define $v_2 = \ln c_{j,t}$, $dv_1 = e^{-\rho t}$. Then $dv_2 = \dot{c}_j / c_j = \gamma = constant$, in steady state, and $v_1 = \Leftrightarrow \frac{1}{\rho}\ e^{-\rho t}$. That implies

$$\begin{aligned} \int_0^\infty \ln c_{j,t}\ e^{-\rho t}\ dt &= \Leftrightarrow \frac{1}{\rho}\left[\ln c_{j,t}\ e^{-\rho t}\right]_0^\infty + \frac{1}{\rho}\int_0^\infty \gamma\ e^{-\rho t}\ dt \\ &= \frac{\ln c_j(0)}{\rho}) \Leftrightarrow \frac{1}{\rho^2}\Big|_0^\infty \gamma\ e^{-\rho t}, \end{aligned} \tag{C1}$$

where $j = 0, 1$. Evaluation of the expression at the particular limits establishes V^h in (13) and V^l in (14).

D Proof that $S(x) > 0$ for all $x \in [0,1]$

I want to show $S(x) \equiv \Delta_2(x) + \Delta_1(x) > 0$ for all $x \in [0,1]$. To this end let $\Delta_3(x) \equiv \alpha(1+x^{\alpha-1})+\rho$ and notice that S is the sum of two functions. I will show that $S(x)$ is *strictly decreasing* in x for any $x \in [0,1]$. So I need to know the derivative of $S(x)$ with respect to x which is given by

$$\begin{aligned}
\frac{\partial S}{\partial x} &= \frac{\partial \Delta_2}{\partial x} + \frac{\partial \Delta_1}{\partial x} \\
&= \Leftrightarrow \frac{\alpha(1 \Leftrightarrow \alpha)x^{\alpha-2}}{\Delta_3} + \frac{\alpha(1 \Leftrightarrow \alpha)^2 x^{\alpha-2}}{\Delta_3} \Leftrightarrow \frac{\alpha^2(1 \Leftrightarrow \alpha)^2 x^{2\alpha-3}}{\Delta_3^2} \\
&= \Leftrightarrow \frac{\alpha^2(1 \Leftrightarrow \alpha)^2 x^{\alpha-2}}{\Delta_3} \Leftrightarrow \frac{\alpha^2(1 \Leftrightarrow \alpha)^2 x^{2\alpha-3}}{\Delta_3^2}
\end{aligned}$$

and is clearly negative for any non-negative x. Thus, $S(x)$ is strictly decreasing in x. Next, I show that $\lim_{x\to\infty} S(x) = 0$ implying $\inf S(x) = 0, \forall x \in (0,\infty)$. So I need to show that $\lim_{x\to\infty} \Delta_2 + \lim_{x\to\infty} \Delta_1 = 0$. Notice that

$$\Delta_1 = \Leftrightarrow \frac{\alpha(1 \Leftrightarrow \alpha)x^{\alpha-1}}{\alpha(1+x^{\alpha-1})+\rho} = \Leftrightarrow \frac{(1 \Leftrightarrow \alpha)x^{\alpha-1}}{1+x^{\alpha-1}+\frac{\rho}{\alpha}} = \Leftrightarrow \frac{(1 \Leftrightarrow \alpha)}{x^{1-\alpha}(1+\frac{\rho}{\alpha})+1}.$$

Then the claim is clearly true since

$$\begin{aligned}
\lim_{x\to\infty} \Delta_1 &= \lim_{x\to\infty} \left(\Leftrightarrow \frac{(1 \Leftrightarrow \alpha)}{x^{1-\alpha}(1+\frac{\rho}{\alpha})+1} \right) = 0 \quad \text{and} \\
\lim_{x\to\infty} \Delta_2 &= \lim_{x\to\infty} \ln \left(\frac{\alpha(1+x^{\alpha-1})+\rho}{\alpha+\rho} \right) = 0.
\end{aligned}$$

Since $x \in [0,1]$ one clearly has $x < \infty$ and so $\inf S(x) > 0$ for all $x \in [0,1]$ and so $S(x) > 0$.

E Proof that $W^u(\hat{x}) > W^u(1)$

I want to show that the level of welfare $W(x)$ is lower at $x = 1$ than at $\hat{x}$. Let $\check{x} = 1$. Then the difference in welfare levels is given by

$$W^u(\hat{x}) \Leftrightarrow W^u(1) = \frac{\hat{x}\ln\hat{c}_1 \Leftrightarrow \ln\check{c}_1}{\rho} + \frac{(1 \Leftrightarrow \hat{x})\ln c_0 \Leftrightarrow (1 \Leftrightarrow \check{x})\ln c_0}{\rho} + \frac{\hat{\gamma} \Leftrightarrow \check{\gamma}}{\rho^2}.$$

where c_0 is independent of x. Furthermore, using (9),

$$\begin{aligned}
\hat{\gamma} \Leftrightarrow \check{\gamma} &= (1 \Leftrightarrow \alpha)\left[1 + \hat{x}^{\alpha}\right] \Leftrightarrow \hat{x}^{\frac{1}{\epsilon}} \Leftrightarrow \rho \Leftrightarrow [(1 \Leftrightarrow a)2 \Leftrightarrow 1 \Leftrightarrow \rho] = \alpha + (1 \Leftrightarrow \alpha)\hat{x}^{\alpha} \Leftrightarrow \hat{x}^{\frac{1}{\epsilon}} \\
&= \alpha + \hat{x}^{\alpha}\left[(1 \Leftrightarrow \alpha)(1 \Leftrightarrow \epsilon\alpha)\right] \equiv B > 0.
\end{aligned}$$

Then the condition for the difference in welfare levels to be positive is

$$\frac{\hat{x}\ln\hat{c}_1 \Leftrightarrow \check{x}\ln\check{c}_1}{\rho} + \frac{(\check{x} \Leftrightarrow \hat{x})\ln c_0}{\rho} + \frac{B}{\rho^2} > 0 \quad \Leftrightarrow \quad \left(\frac{\hat{c}_1}{c_0}\right)^{d} e^{\frac{B}{\rho\check{x}}} > \frac{\check{c}_1}{c_0}$$

where $d = \frac{\hat{x}}{\check{x}}$ and $\check{x} = 1$. Note that $\left(\frac{\hat{c}_1}{c_0}\right)^{d} > 1$ and $e^{y} = \left(1 + \frac{y}{1!} + \frac{y^2}{2!} + \frac{y^3}{3!} + \ldots\right)$. Then a sufficient condition for the inequality with $\check{x} = 1$ to hold is

$$\begin{aligned}
\left(1 + \frac{B}{\check{x}\rho} + \ldots\right) &> 1 + \frac{\alpha\check{x}^{\alpha-1}}{\alpha + \rho} \\
C + B &> \frac{\rho\alpha}{\alpha + \rho} \\
C + \hat{x}^{\alpha}\left[(1 \Leftrightarrow \alpha)(1 \Leftrightarrow \epsilon\alpha)\right] &> \frac{\rho\alpha}{\alpha + \rho} \Leftrightarrow \alpha \qquad \text{(E1)}
\end{aligned}$$

where C is a positive constant. Thus, the inequality holds because the RHS is negative. Hence, $W^u(\hat{x}) > W^u(1)$ implying that $v(1) < 0$, that is, marginal welfare is negative if x is close to one.[20]

[20]It is not difficult to verify that a similar result may be obtained for $\tilde{x} > \hat{x}$. If $\gamma - \tilde{\gamma} > \rho$ then $W^u(\hat{x}) > W^u(\tilde{x})$ and $v(\tilde{x}) < 0$ by reasoning as above.

References

Aghion, P. and P. Howitt, *Endogenous Growth Theory*, Cambridge, Mass.: MIT Press, 1998.

Alesina, A. and D. Rodrik, "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 1994, *109*, 465–490.

Atkinson, A. B., "On the Measurement of Inequality," *Journal of Economic Theory*, 1970, *2*, 244–263.

Azariadis, C. and A. Drazen, "Threshold Externalities in Economic Development," *Quarterly Journal of Economics*, 1990, *105*, 501–526.

Becker, G. S, K. M. Murphy, and R. Tamura, "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy*, 1990, *98*, S12–S37.

Bénabou, R., "Inequality and Growth," in B. S. Bernanke and J. J. Rotemberg, eds., *NBER Macroeconomics Annual 1996*, Cambridge and London: MIT Press, 1996, pp. 11–73.

Bertola, G., "Factor Shares and Savings in Endogenous Growth," *American Economic Review*, 1993, *83*, 1184–1198.

——, "Macroeconomics of Distribution and Growth," Working Paper 98/39, European University Institute 1998.

Bound, J. and G. Johnson, "Changes in the Structure of Wages During the 1980's: An Evaluation of Alternative Explanations," Working Paper 2983, NBER 1989.

Caballé, J. and M. S. Santos, "On Endogenous Growth with Physical and Human Capital," *Journal of Political Economy*, 1993, *101*, 1042–1067.

Chiang, A. C., *Elements of Dynamic Optimization*, New York: McGraw-Hill, 1992.

Chiu, W. H., "Income Inequality, Human Capital Accumulation and Economic Performance," *Economic Journal*, 1998, *108*, 44–59.

Deardorff, A. V., "Growth and International Investment with Diverging Populations," *Oxford Economic Papers*, 1994, *46*, 477–491.

Deininger, K. and L. Squire, "A New Data Set Measuring Income Inequality," *World Bank Economic Review*, 1996, *10*, 565–591.

Eckstein, Z. and I. Zilcha, "The effects of compulsory schooling on growth, income distribution and welfare," *Journal of Public Economics*, 1994, *54*, 339–359.

Fields, G. S., "Measuring inequality change in an economy with income growth," *Journal of Development Economics*, 1987, *91*, 611–625.

Findlay, R. and H. Kierzkowski, "International Trade and Human Capital: A Simple General Equilibrium Model," *Journal of Political Economy*, 1983, *91*, 957–978.

Freeman, R., "The Decline in the Economic Rewards to College," *Review of Economics and Statistics*, 1977, *59*, 18–29.

Garcia-Peñalosa, C., "Absolute Poverty and Relative Inequality: A Theory of Growth and Wage Differentials," Discussion Paper 97, Nuffield College, Oxford, UK 1995.

———, "The Paradox of Education or the Good Side of Inequality," *Oxford Economic Papers*, 1995, *47*, 265–285.

Glomm, G. and B. Ravikumar, "Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality," *Journal of Political Economy*, 1992, *100*, 818–834.

Hirshleifer, J. and J. G. Riley, *The Analytics of Uncertainty and Information*, Cambridge, U.K.: Cambridge University Press, 1992.

Katz, L. F. and A. Revenga, "Changes in the Structure of Wages: The US versus Japan," Working Paper 3021, NBER 1989.

Kremer, M., "Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics*, 1993, *108*, 681–716.

Lambert, P. J., *The Distribution and Redistribution of Income: A Mathematical Analysis*, 2nd ed., Manchester, UK: Manchester University Press, 1993.

Londoño, J. L., "Human Capital and Long Run Swings of Income Distribution," mimeo, Harvard University 1990.

Lucas, R. E., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 1988, *22*, 3–42.

Mas-Colell, A., M. D. Whinston, and J. R. Green, *Microeconomic Theory*, Oxford University Press, 1995.

Mincer, J., "Investment in human capital and personal income distribution," *Journal of Political Economy*, 1958, *66*, 281–302.

Nerlove, M., A. Razin, E. Sadka, and R. K. von Weizsäcker, "Comprehensive income taxation, investments in human and physical capital, and productivity," *Journal of Public Finance*, 1993, *50*, 397–406.

Razin, A. and C.-W. Yuen, "Capital income taxation and long-run growth: New perspectives," *Journal of Public Economics*, 1996, *59*, 239–263.

Rehme, G., "Essays on Distributive Policies and Economic Growth," Unpublished PhD Thesis, European University Institute 1998.

Romer, D., *Advanced Macroeconomics*, New York: McGraw Hill, 1996.

Romer, P. M., "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 1986, *94*, 1002–1037.

Rosenzweig, M. R., "Population Growth and Human Capital Investments: Theory and Evidence," *Journal of Political Economy*, 1990, *98*, S38–S70.

Shorrocks, A. F., "Ranking Income Distributions," *Economica*, 1983, *50*, 1–17.

Tamura, R., "Income Convergence in an Endogenous Growth Model," *Journal of Political Economy*, 1991, *99*, 522–540.

Tilak, J., "Rates of Return to Education and Income Distribution," *De Economist*, 1989, *137*, 454–465.