Reprinted from

JOURNAL OF PUBLIC ECONOMICS

The effects of compulsory schooling on growth, income distribution and welfare

Zvi Eckstein

Department of Economics, Tel-Aviv University, Ramat Aviv, Tel-Aviv 69978, Israel and Boston University, Boston, MA 02215, USA

Itzhak Zilcha*

Department of Economics, Tel-Aviv University, Ramat Aviv, Tel-Aviv 69978, Israel

N·H

ELSEVIER

The effects of compulsory schooling on growth, income distribution and welfare

Zvi Eckstein

Department of Economics, Tel-Aviv University, Ramat Aviv, Tel-Aviv 69978, Israel and Boston University, Boston, MA 02215, USA

Itzhak Zilcha*

Department of Economics, Tel-Aviv University, Ramat Aviv, Tel-Aviv 69978, Israel

In this paper we consider an OLG model with productive capital and human capital affecting the quality of labor. In each generation parents invest in their children's education but disregard its external effect on the aggregate production function. Government intervention in providing compulsory schooling increases economic growth (along the equilibrium path), while the intragenerational income distribution becomes more equal. Also, in the long run, the majority of individuals in each generation are better off due to compulsory schooling.

1. Introduction

Compulsory basic education is taken for granted in modern societies. The principle of placing all children in school for a specified period of time, which began in the eighteenth and nineteenth centuries, was later adopted by European governments.[1] During this century and, in particular, after the Second World War all countries adopted the practice of compulsory education [OECD (1983)]. The exact form of compulsory education varies across countries, but despite differences in most other areas there is a unanimous acceptance of this policy as an essential basic public service. The compulsory education law is widely observed and enforced in all countries,

[1]Frederic the Great of Prussia and Maria Theresa of Austria started compulsory elementary education as early as 1763 and 1773, resepctively [Melton (1988)]. In England, compulsory education was imposed by the Elementary Education Act of 1870 [West (1970)].

and the main part of public expenditures for education is devoted to the institutions providing education for children.

Why is it that compulsory education is widely accepted as an important basic public service? Why is most of the public support for education provided by this institution and not by a tax-transfer method even in economies where intervention by the state avoids compulsory rules?[2]

All the classical economists [Smith, Malthus, Bentham and J.S. Mill; see West (1970, pp. 111-112)] claimed that an increase in the education level of the poor would result in a decrease in crime and disorder (negative externality)[3] and since the state imposes the law and order, a cost-benefit calculation can justify the introduction of compulsory education. The classical economists realized the economic gains from education but thought that the division of labor in the free market would internalize all the economic benefits.

As pointed out by J.S. Mill, in the case of children's education the principle of self-interest breaks down since 'the person most interested is not the best judge of the matter', and it is not clear that parents make the best judgment for their children [West (1970, pp. 7-11)]. Mill, however, thought that 'a general state education is ... moulding people to be exactly like one another', and given his support for variety and liberty he rejected the idea of compulsory state education [West (1970, p. 124)].

However, the argument that the state should protect children from parents who do not value education has been viewed as a sufficiently negative externality to justify the law of compulsory minimum education. Educators emphasize the principle of 'equal opportunities' for children, which seems to us to stem from both the need to protect children who otherwise would have received a very low level of education, and the social goal of a more equal distribution of income.

Already in the early years of the nineteenth century it was perceived by McCulloch in his *Principles of Political Economy* (1825) that 'A better system of education and a better law of inheritance are the two most powerful means of reducing inequalities of income' [see Dalton (1920, pp. 56-58)]. This work provides some support for this view.

In this paper we consider an economy where heterogeneity among individuals in each generation is a consequence of the different preferences of parents with respect to the education level of their children. Parents' preferences depend on their child's level of education, or human capital, but

---

[2]Angrist and Krueger (1990, 1991) have found that compulsory schooling in the United States is effective so that a significant larger fraction of students were constrained to enter school early and dropout was delayed such that the duration of school attendance was increased due to compulsory school laws.

[3]Friedman (1962) emphasized the need for a minimum level of education for the better functioning of a democratic society.

are independent of the child's income or lifetime utility. On the other hand, the level of education obtained by each child has an important impact on his ability as a worker and in the continuation of the learning process [Becker (1975)]. As a result, when parents are homogeneous (as in section 2), the education level is less than the optimal level. The first-best policy which corrects this negative externality constitutes intergenerational taxes and transfers which would guarantee the optimal investment in children's education.[4]

To guarantee non-degenerate income distribution we assume that parents' preferences regarding their offspring's education are different;[5] hence, investment in their children's human capital differ. The only way to identify an individual is by observing his investment in his child's human capital. As a result, the implementation of such a policy would be 'too late'; furthermore, the actual tax-transfer program should be implemented specifically for each individual at each date. Therefore, such a policy cannot be considered seriously (high implementation costs) in an economy where indiviuduals are heterogeneous with respect to their preferences and their earnings (human capital).

Compulsory education can be viewed as a potential second-best alternative. The fact that this particular method of intervention has been implemented in all countries independent of other aspects, such as the political environment, implies that it should combine several positive elements that are widely shared by people. To investigate the economic aspects of this observation we compare the dynamic allocation of our overlapping generations model without intervention with a model where compulsory elementary schooling is financed by a proportional tax on wage income.

We show that a certain minimum level of compulsory schooling, financed by a proportional tax rate on wage income, increases the aggregate output

---

[4]This policy is equivalent to a model that assumes that parents' utility function is an increasing function of the children's income [Becker and Tomes (1979) and Saint-Paul and Verdier (1991)]. We assume that preferences of the current population depend only on variables that are directly influenced by the parents' decisions so that full neutrality of future policies is not attained (Bernheim and Bagwell (1988)].

[5]Heterogeneity among parents' investment in their children's education is a major source of the observed income distribution [Becker and Tomes (1979) and Loury (1981)]. In particular, optimal investment in children's education requires that parents take into consideration the income gain to their children due to their education and the ability of at least some parents to take loans in order to educate their children. The inability of the market to implement a contract in which children pay back the loans from their parents imply an underinvestment in human capital. A possible policy that seems natural for such an environment is that the government would enforce such contracts. Is such a policy less reasonable than a compulsory education? Note that we could introduce heterogeneity by assuming that only part of the population is not fully altruistic (Aiyagari (1989)]. Hence, only a few, randomly chosen parents do not consider the marginal income gain to their children. As a result, the current investment in the education of these parents implies an important source of negative externality on the level of investment in human capital.

and, at the same time, reduces the level of inequality in income distribution. In his empirical work Chiswick (1969) argues that minimum schooling laws increase the level of skewness but decrease the inequality in the distribution of income. We also prove that in the long run, the majority of individuals in each generation are better off under some level of compulsory education.[6] We show that compulsory education induces more investment in schooling and 'improves' the human capital distribution. As a result, after several periods, the output level increases at each subsequent date vis-à-vis an economy without this policy. Income distribution, starting at $t = 0$, becomes more equal since the program involves a transfer of resources from the rich to the poor and eliminates the frequency of a population with very low investment in human capital. The median individual is better off since his/her income is affected by the equality aspects of this policy and the economy-wide additional growth. It seems to us that the externality embedded in parents' decisions combined with the heterogeneity of parents' preferences regarding the quality (human capital) of their children, provide the main argument for compulsory elementary education.

Recently we have witnessed a renewed interest in income distribution, growth and the investment in human capital. Lucas (1988) and Azariadis and Drazen (1990) analyzed the role of the investment in human capital and externalities on the long-run growth rate of the economy. We have adopted a similar function for the accumulation of human capital but we emphasize the role of the parents.[7] Saint-Paul and Verdier (1991) analyze the relationship between public education, growth and income distribution in a model which is related to ours, but has different features and motivation. Persson and Tabellini (1991) provide some evidence that growth rates are positively associated with more equality in income in a cross-section of nations. This evidence is consistent with their model of income distribution and growth as well as with most other recent papers on this subject.

The paper is organized as follows. In the next sections we discuss a benchmark model with a representative agent in order to demonstrate the inefficiency of the equilibrium allocation. In section 3 we present the model with heterogeneous population. The positive implications of compulsory education are discussed in section 4 and welfare implications in section 5. Section 6 concludes the paper and the appendix contains the proofs.

2. A benchmark model

Individuals live for two periods in an OLG economy as in Diamond

---

[6]Loury (1981) and Saint-Paul and Verdier (1991) have similar results using different models with altruistic parents.

[7]Perotti (1991) and Fernandez and Rogerson (1991) analyze the role of tax subsidies to education through a majority voting model in the determination of income distribution and growth.

(1965). We assume that each individual works during her first period ('young') and only consumes in her second period, i.e. when 'old'. During the first period she gives birth to $1+n$ offsprings and thus has to allocate some of her time during this period (where endowment of time is a constant) to educating them. This has a direct effect upon her children's human capital, or 'knowledge', which is assumed to be a source of utility to the parent. The lifetime utility function of $i \in G_t$ (the ith generation, i.e. all individuals born at date t) is $u_i \colon R_+^4 \to R^1$, $u_i(c_t^{yi}, c_{t+1}^{oi}, z_t^i, (1+n)h_{t+1}^i)$, where c_t^{yi} and c_{t+1}^{oi} are the consumption when young (date t) and when old (date $t+1$), z_t^i represents the leisure, h_{t+1}^i represents the human capital of i's offspring. Parents can affect the human capital of their offspring only by the time devoted in educating this child, hence we disregard random factors. Each utility function u is strictly concave, increasing, continuously differentiable and $(\partial u/\partial c_j)(x) = \infty$ if $x_j = 0$, $j = 1, 2$. The evolution of human capital process over time depends upon the parent's level of human capital and the effort (measured in time) invested by them in raising their offspring. We assume that for some $\frac{1}{2} < \beta \leqq 1$, the evolution of human capital is according to

$$h_{t+1}^i = A(e_t^i)[h_t^i]^\beta, \tag{1}$$

where e_t^i is the parents' investment (in time) in education,[8] where the function $A(\cdot)$ satisfies $0 < A(0) \leqq 1$, $A(1) > 1$, $A' > 0$, $A'' \leqq 0$. We assume that the rate of population growth is n, $n \geqq 0$, and that the set of individuals in generation 0 is given by $[0, 1]$.

The aggregate level of human capital at each date t has a direct effect upon the production possibilities at that period. In particular we take

$$Y_t = F(K_t, L_t) \tag{2}$$

to be the aggregate production function, where $L_t = l_t h_t$ is the effective aggregate labor and K_t the aggregate capital stock. $F(\cdot, \cdot)$ is assumed to exhibit constant returns to scale, it is strictly increasing, concave, continuously differentiable and satisfies $F_1(0, L) = \infty$, $F_2(K, 0) = \infty$, $F(0, L) = F(K, 0) = 0$, and $-LF_{LL}/F_L \leqq 1$.[9]

Our model basically provides some type of Harrod neutral technological progress due to accumulation in knowledge or human capital. We assume that for each individual $l_t = 1$, that is, labor supply is inelastic.[10] Production at date t takes place by competitive firms who borrow capital at date $t-1$

---

[8]This specification is similar to that of Loury (1981) and Azariadis and Drazen (1990).

[9]A *sufficient* condition which guarantees this property about the elasticity of F_L is that the production function's elasticity of substitution is greater than or equal to 1.

[10]This assumption simplifies the model. A further simplification would be to ignore the accumulation of physical capital to emphasize the role of human capital. On the other hand, it is possible to extend the model by introducing increasing returns to scale due to external effects of human and physical capital [see, for example, Romer (1986) and Lucas (1988)].

and hire labor services at date t. The factor prices are given by the marginal product. Since the human capital of a worker is observable, the wage payments will depend upon the effective labor supply of the worker, i.e. $W_t^i = w_t h_t^i$, where $w_t = F_2(K_t, L_t)$ is the wage rate. K_{t+1} os the aggregate savings at date t and the competitive interest rate is $R_{t+1} = F_1(K_{t+1}, L_{t+1})$.

Let us consider first an economy with homogeneous population. The optimal choice of the 'young' in G_t is derived from maximizing

$$\max_{s_t, e_t} u(c_t^y, c_{t+1}^o, z_t, (1+n)h_{t+1}) \tag{3}$$

subject to

$$c_t^y = w_t h_t - s_t, \tag{4}$$

$$c_{t+1}^o = s_t R_{t+1}, \tag{5}$$

$$z_t = 1 - (1+n)e_t, \tag{6}$$

$$h_{t+1} = A(e_t)h_t^\beta. \tag{7}$$

After substituting the constraints the first-order conditions with respect to s_t and e_t are

$$-u_1(x_t) + R_{t+1}u_2(x_t) = 0, \tag{8}$$

$$-u_3(x_t) + u_4(x_t)A'(e_t)h_t^\beta \leqq 0,$$

$$= 0 \quad \text{if} \quad e_t > 0, \tag{9}$$

where $x_t = [c_t^y, c_{t+1}^o, z_t, (1+n)h_{t+1}]$.

It is clear that the optimal amount of time invested in educating the children disregards the gain by the child from this investment. This is a result of our assumption that neither the child's income not his utility enter the parent's objective function. Hence, the wage increase in the next period due to an additional investment in human capital does not affect the parents' allocation of time. Note that it is possible that at the optimum $e_t = 0$ for all t, hence if $\beta < 1$, h_t decreases over time. However, if u_4 at $h_{t+1} = 0$ is large, then human capital (and output) will be bounded away from 0.

In order to analyze the efficiency of the competitive allocations attained in the above economy we shall consider the case of a planner (or a dynastic model) where the problem at $t = 0$ is to maximize the present value of a discounted stream of future utilities. Consider some positive monotone decreasing sequence (λ_t), $\sum_{t=0}^{\infty} \lambda_t < \infty$. Given k_0 and h_0 the objective of the planner is to maximize

$$\max_{\{c_t^0, h_{t+1}, k_{t+1}\}} \sum_{t=0}^{\infty} \lambda_t u(x_t) \tag{10}$$

s.t.

$$c_t^y + \frac{c_t^0}{1+n} + (1+n)k_{t+1} = F(k_t, h_t), \tag{11}$$

$$z_t = 1 - (1+n)e_t \geq 0, \tag{12}$$

$$h_{t+1} = A(e_t)h_t^\beta, \tag{13}$$

for $t = 0, 1, 2, \ldots$.

By inserting constraints (11)–(13) into (10), substituting e_t in x_t by $A^{-1}(h_{t+1}/h_t^\beta)$ and differentiating with respect to k_{t+1}, c_t^0 and h_{t+1} we obtain the first-order conditions for the planner's problem:

$$\lambda_{t+1} u_1(x_{t+1}) F_1(k_{t+1}, (1+n)h_{t+1}) - \lambda_t(1+n)u_1(x_t) = 0, \tag{14}$$

$$\lambda_t u_2(x_t) - \lambda_{t+1} u_1(x_{t+1})/1+n = 0, \tag{15}$$

$$u_3(x_t) = h_t^\beta A'(e_t)u_4(x_t) + (1+n)u_2(x_t)h_t^\beta A'(e_t)F_2(k_{t+1}, (1+n)h_{t+1})$$

$$+ \beta(1+n)u_2(x_t)\frac{u_3(x_{t+1})}{u_1(x_{t+1})} \frac{A'(e_t)}{A'(e_{t+1})} h_t^\beta h_{t+1}^{-\beta-1} h_{t+2}. \tag{16}$$

From (14) and (15) we derive that

$$u_1(x_t)/u_2(x_t) = F_1(k_{t+1}, (1+n)h_{t+1}). \tag{17}$$

Now we are ready to show that the competitive allocation can be dominated by such a plan; hence, it is not optimal since parents under-invest in the human capital of their children. The reason for that is that they under-value the benefits of this investment by ignoring its impact upon the wage and leisure of their offspring.

Comparing eqs. (8) and (9) with (16) and (17) one easily sees that (17) coincides with (8) if $R_{t+1} = F_1(k_{t+1}, (1+n)h_{t+1})$. However, eq. (16) includes two additional terms: the first includes F_2 at date $t+1$, reflecting the impact of the parents' decision on the children's earnings; while the second terms (u_3 at date $t+1$) reflects the impact of the parents' investment on the children's tradeoff between leisure and investment in human capital (of their children). Let us denote the competitive allocation by an asterisk, and let the planner's

optimum be denoted by a bar. The following proposition establishes the under-investment in human capital in the competitive economy.

Proposition 1.   Given k_0 and h_0 the competitive equilibrium allocation is not Pareto optimal. Moreover, assuming that u is homothetic and that $u_{43} \geqq 0$ imply that there exists some Pareto-optimal allocation (from k_0 and h_0) with $(\bar{h}_t)$ such that $h_t^* < \bar{h}_t$ for all $t \geqq 1$.

Proof. For the competitive economy we derived condition (17) as in the optimal allocation case, while condition (9) can be written as (assuming e_t^* is positive for all t)

$$u_3(x_t^*)/u_4(x_t^*) = A'(e_t^*)h_t^{*\beta}. \tag{18}$$

Homotheticity of u implies that $u_3(x_t)/u_4(x_t)$ does not depend explicitly on (c_t^y, c_{t+1}^0).

Comparing (18) and (17) with conditions (16) and (17) we conclude that the two paths cannot coincide for any choice of (λ_t). Now, assuming that $u_{43} \geqq 0$ we can show that $u_3(x_t)/u_4(x_t)$ is increasing in e_t. Given k_0 and h_0 we find from (16) and (18) [since the right-hand side in (16) is larger than the right-hand side in (18) at $t = 0$] that $e_0^* < \bar{e}_0$. Therefore, $h_1^* < \bar{h}_1$ and hence, by (16) and (18) for $t = 1$, we obtain that $e_1^* < \bar{e}_1$. This process can be continued for $t = 2, 3, \ldots$.   □

Given the separability, or complementarity, between leisure and the human capital of children, the economy without any government intervention is characterized by under-investment in human capital. To construct a policy that internalizes the externality in investments in human capital in this competitive economy, one has to construct a rule where the parents' decision is affected by the wage earnings and leisure of the children. Note that the inclusion of the children's income in the parents' utility [e.g. Becker and Tomes (1979)] does not guarantee an optimal allocation.

It seems to us completely unreasonable to assume that there might be an institution that would transfer income between children and parents in the way required by such an optimal policy. If we assume that parents have a bequest motive, such as in Barro (1974), then each parent solves an infinite horizon social planning problem, but we encounter in this case the same problems raised by Bernheim and Bagwell (1988).

However, suppose that parents are heterogeneous with respect to their preferences on the quality of their children. Furthermore, suppose that the attitude of an individual towards investment in his child's human capital is a random variable which is known to the agent when he is young but was not known to his parents. In that case the tax-transfer of the optimal policy should be individual-specific and since preferences are not observable, there

is no way to implement such a policy. Moreover, it is unclear how reasonable it is to assume that each parent solves the dynamic optimization [see Bernheim and Bagwell (1988)] particularly when the preferences of the coming generations are unknown.

3. Heterogeneous population

Following the above discussion we introduce heterogeneity into our economy by assuming that each agent's taste for human capital of her child is a random draw from an independent process. That is, in each generation t individuals are alike except in the intensity of their utility from the human capital (of their offspring) h_{t+1}.

We assume that each generation G_t has a continuum of individuals, say the interval $[0, 1]$; thus we assume now that there is *no population growth* in this economy. The utility function of each individual is determined at the outset of his lifetime by some random process. These random variables will be independent and identically distributed in each generation and across generations. To state this more precisely, let us denote by $\theta \in [0, 1]$ a 'dynasty', i.e. an infinite sequence of individuals related to each other as a family (i.e. 'parent' and 'child'). Let $\tilde{y}$ be a random variable with a given distribution on $[a, b]$, $0 < a < b < \infty$. For each $i \in G_t$ who belongs to the family (or dynasty) named θ, $\theta \in [0, 1]$, there corresponds a random variable $\tilde{\omega}_t^\theta$ distributed as $\tilde{y}$. Moreover, these random variables are i.i.d. with respect to t and θ.[11] The realization of $\tilde{\omega}_t^\theta$ will affect each individual's taste regarding the choice between leisure and the human capital of his offspring.

Denote by $\Omega_t = \{\omega^t = (\omega_0, \dots, \omega_t) \mid \omega_k \in [a, b] \text{ for } 0 \leq k \leq t\}$, i.e. the set of all possible histories at date t. For each θ in G_{t+1} her human capital level depends upon the history of his family, θ, up to date t, i.e. on $\omega^{\theta t} = (\omega_0^\theta, \omega_1^\theta, \dots, \omega_t^\theta) \in \Omega_t$. Given the above stochastic process, $\tilde{y}$, there exist a probability distribution function, μ_t (defined upon the Borel sets in Ω_t), which describes the distribution of $(\omega^{\theta t})_{\theta \in G_t}$. For each $\theta \in G_t$ the consumption at date t, c_t, for example, is a function of $\omega^{\theta t}$, i.e. $c_t = c_t(\omega^{\theta t})$. Notice that ω_t^θ is revealed at the outset of date t, and hence each individual knows her utility function when she makes her decisions about consumption, leisure, and investment in his child's human capital. The utility of $\theta \in G_t$ is given by (note that $z_t = 1 - e_t$)

$$U_t^\theta = (c_t^y)^{\alpha_1} (c_{t+1}^0)^{\alpha_2} (z_t)^{\alpha_3} (h_{t+1})^{\alpha_4(\omega_t^\theta)},$$

where α_4 depends on the realization of $\tilde{\omega}_t^\theta$. Moreover, as we have assumed,

$$h_{t+1}(\omega^{\theta t}) = A(e_t(\omega^{\theta t}))[h_t(\omega^{\theta(t-1)})]^\beta, \quad t = 0, 1, \dots, \tag{19}$$

---

[11]One should be careful in making such an assumption since there is a continuum of families in each generation; see Judd (1985).

where $\frac{1}{2} < \beta \leq 1$ and $e_t(\omega^{\theta t})$ is the (time) investment in educating the offspring. We assume that $\alpha_4(\cdot)$ is a continuous and increasing function on $[a, b]$. Since labor supply is inelastic, the aggregate effective labor at each date t is given by

$$L_{t+1} = \int h_{t+1}(\omega^{\theta t}) \, d\mu_t(\omega^{\theta t}). \tag{20}$$

The human capital of a worker is observable, hence the wage payment will depend upon the effective labor supply, i.e. $W_t^i = w_t h_t^i$, where $w_t = F_2(K_t, L_t)$ is the wage rate. To simplify our notation we will write $c_t(\omega)$ instead of $c_t(\omega^{\theta t})$; also the integral $\int c_t(\omega^{\theta t}) \, d\mu_t(\omega^{\theta t})$ will be denoted by $\int c_t(\omega) \, d\mu_t$.

We assume that the government provides compulsory schooling (CS) financed by taxes on income in the following manner. In each period we take the human capital of publicly provided education to be the *average* human capital of the population at that period, denoted by H_t^g; thus $H_t^g = \int h_t(\omega) \, d\mu_{t-1}$.[12] The level of this compulsory education (provided to all young members of generation t) at period t is denoted by e_t^g. Now we take the evolution of the human capital process to be given by[13]

$$h_{t+1}(\omega) = A(e_t^g + e_t(\omega))(\bar{h}_t(\omega))^\beta, \tag{21a}$$

where $\bar{h}_t$ is the 'relevant' human capital level which affects h_{t+1}. We choose $\bar{h}_t$ to be the weighted average of the human capital at the public schooling and the parents' human capital,

$$\bar{h}_t(\omega) = \frac{e_t^g H_t^g + e_t(\omega) h_t(\omega)}{e_t^g + e_t(\omega)}. \tag{21b}$$

The compulsory education is financed by proportional taxes on wage income and we denote by τ_t the tax rate at date t. Each $\theta \in G_t$ pays, given the wage rate w_t at date t, $T_t = \tau_t w_t h_t(\omega^\theta)$. Thus, given w_t and R_{t+1} the interest on savings, the tax rate τ_t, e_t^g and H_t^g, each individual maximizes her lifetime utility function under these conditions. That is, she chooses saving, s_t, and additional time invested in educating her own offspring, e_t, such that she solves the following problem:

$$\max_{s_t, e_t} [w_t h_t(\omega)(1 - \tau_t) - s_t]^{\alpha_1} [s_t R_{t+1}]^{\alpha_2} [1 - e_t]^{\alpha_3} [A(e_t^g + e_t)\bar{h}_t^\beta(\omega)]^{\alpha_4(\omega_t^\theta)}. \tag{22}$$

---

[12]Loury (1981) assumes that public education provides the average investment in human capital of a CE. One could consider alternative assumptions on the quality of public education. Note that changes in quality affect taxes and/or the total level of compulsory schooling.

[13]The specification of the production of human capital with CS depends on four factors of production. Our choice of (21a) and (21b) seems a reasonable one. Of course, one can choose another process where $\bar{h}_t$ is some other function of H_t^g and $h_t(\omega)$. Each particular specification might have some implications on the results.

Necessary and sufficient conditions for an optimum are

$$\frac{s_t(\omega^\theta)}{w_t h_t(\omega^\theta)(1-\tau_t)-s_t(\omega^\theta)} = \frac{\alpha_2}{\alpha_1}, \tag{23}$$

$$\frac{A(e_t^g + e_t(\omega^\theta))[\bar{h}_t(\omega^\theta)]^\beta}{1 - e_t(\omega^\theta)} \gtreqless \frac{\alpha_4(\omega_t^\theta)}{\alpha_3} A'(\bar{h}_t)^\beta + \beta A(\bar{h}_t')[\bar{h}_t(\omega)]^{\beta-1}, \tag{24}$$

with equality in (24) whenever $e_t(\omega^\theta) > 0$. Denote the optimum by an asterisk; hence,

$$c_t^{y*}(\omega) = w_t h_t(\omega)(1-\tau_t) - s_t^*(\omega), \tag{25}$$

$$c_{t+1}^{o*}(\omega) = s_t^*(\omega) R_{t+1}, \tag{26}$$

$$z_t^*(\omega) = 1 - e_t^*(\omega), \tag{27}$$

$$h_{t+1}^*(\omega) = A(e_t^g + e_t^*(\omega))[\bar{h}_t(\omega)]^\beta. \tag{28}$$

We shall consider compulsory education plans which satisfy the following two properties: (a) the human capital level of the educators is the average of the population for that generation, and (b) the compulsory education is fully financed by the taxes at each date. Namely, for each period t, the expenditure should equal the total amount of taxes collected:

$$w_t e_t^g H_t^g = \int \tau_t w_t h_t(\omega) \, d\mu_{t-1},$$

which implies that $e_t^g = \tau_t$ for all t.[14] As a result, the effective labor supply (i.e. the labor supply used in the production process) with $\tau_t > 0$ is given by $L_t = (1-\tau_t) \int h_t^*(\omega) \, d\mu_{t-1}$.

Given the initial capital stock, K_0, the human capital distribution at period 0, $h_0(\omega)$, and the tax rates, τ_t, to finance compulsory education, a *competitive equilibrium* (CE) is a $\langle c_0^{o*}(\omega), (c_t^{y*}(\omega), c_{t+1}^{o*}(\omega), e_t^*(\omega))_{t=0}^\infty, (w_t, R_{t+1})_{t=0}^\infty \rangle$ that satisfies the following conditions:

(a) $(c_t^{y*}(\omega), c_{t+1}^{o*}(\omega), e_t^*(\omega))$ is the optimum for (22) for all ω, for $t = 0, 1, \ldots$,

(b) $L_t^* = (1-\tau_t) \int h_t^*(\omega) \, d\mu_{t-1}$ and $K_t^* = \int s_{t-1}^*(\omega) \, d\mu_{t-1}$, for $t = 1, 2, \ldots$,

(c) $w_t = F_2(K_t^*, L_t^*)$, for $t = 0, 1, 2, \ldots$, $\qquad\qquad\qquad$ (29)

(d) $R_{t+1} = F_1(K_{t+1}^*, L_{t+1}^*)$, for $t = 1, 2, \ldots$ $\qquad\qquad$ (30)

(e) $w_t e_t^g H_t^g = \tau_t w_t \int h_t^*(\omega) \, d\mu_{t-1}$, for $t = 0, 1, \ldots$. $\qquad$ (31)

---

[14]Note that if τ_t depends on ω (e.g. progressive taxes), then the level of compulsory education would not be equal to the tax rate.

Thus the effective wages are the marginal product of the effective labor, L_t^*, i.e. the effective labor applied in the production process (not including efforts used to raise the quality of labor through education, $e_t^g H_t^g$). Interest factors are the marginal product of capital, K_t^*. Condition (31) guarantees that the cost of compulsory education is covered by the taxes collected at equilibrium and it is easy to show that the competitive equilibrium satisfies the material balance conditions:[15]

$$\int c_t^{y*}(\omega)\,d\mu_t + \int c_t^{0*}(\omega)\,d\mu_t + K_{t+1}^* = F(K_t^*, L_t^*), \quad t = 0, 1, \ldots \tag{32}$$

Without loss of generality we shall assume that $\tau_t = \tau$ for all t and that the equilibrium $\{K_t^*/L_t^*\}_{t=0}^{\infty}$ is bounded.

4. Growth and distribution

We first examine the effects of compulsory schooling on the rate of growth in equilibrium. Specifically, we compare the CE when $\tau = 0$ with the CE when $\tau > 0$ (not too large), i.e. the case where a certain positive level of compulsory education is imposed.

Proposition 2. Suppose that, $eA'(e)/A(e)$ is non-increasing. Given K_0 and $h_0(\omega)$, let $\langle c_0^{0*}(\omega), (c_t^*, c_{t+1}^{0*}, e_t^*)_{t=0}^{\infty}, (w_t, R_{t+1})_{t=0}^{\infty} \rangle$ and $\langle c_0^{0'}, (c_t^{y'}, c_{t+1}^{0'}, e_t')_{t=0}^{\infty}, (w_t', R_{t+1}')_{t=0}^{\infty} \rangle$ be the competitive equilibria with $e^g = 0$ and $e^g > 0$ correspondingly. If e^g is not too large, there exists $N(\tau) < \infty$ such that for all $t \geq N(\tau)$, $K_t' > K_t^*$ and $L_t' > L_t^*$. Moreover, $N(\tau) \to 1$ as $\tau \to 0$.

We relegate all the proofs to the appendix.

This proposition implies that a small level of compulsory education results in higher levels of output, capital and aggregate human capital beginning at some finite date. This is true from period 1 on only for a very low level of compulsory education; in fact, there exists a tradeoff between the level of compulsory schooling and the time interval until growth becomes higher. The main reason for the absence of an immediate increase (i.e. at $t = 1$) in output due to CS is that this policy induces higher investment in education. The reduction in savings due to the new tax lower the capital stock. If taxes are large there is no increase in output, capital and labor. Therefore, assuming that a steady state exists, there is a positive level of τ (or e^g) that maximizes the *steady-state* level of output. The next question is whether higher growth due to compulsory education implies more equality in income distribution.

To study the distributional effect of compulsory education we need a

[15]As we have seen in the homogeneous population case, we cannot expect the CE to be efficient.

formal measure of income inequality. The measure we use here has been introduced by Atkinson (1970) and characterized later by Rothschild and Stiglitz (1973). Given two income distributions, $X(\omega)$ and $Y(\omega)$, with the same mean, denote by $s(\alpha, X)$ the share of total income received by the poorest α percentage of the population. As α varies in $[0,1]$ $s(\alpha, X)$ traces the Lorenz curve associated with X. We say that X is *more equal* (income distribution) than Y if $s(\alpha, X) \geqq s(\alpha, Y)$, for all $\alpha \in [0,1]$ with strict inequality for some α. As was shown by Atkinson and by Rothschild–Stiglitz this is equivalent to a second-degree stochastic dominance (SDSD), i.e. $X >_2 Y$.

Proposition 3. Let $\langle c_0^{0*}, (c_t^{y*}, c_{t+1}^{0*}, e_t^*)_{t=0}^{\infty}, (w_t^*, R_{t+1}^*)_{t=0}^{\infty} \rangle$ *and* $\langle c_0^{0'}, (c_t^{y'}, c_{t+1}^{0'}, e_t')_{t=1}^{\infty}, (w_t', R_{t+1}')_{t=0}^{\infty} \rangle$ *be the CE with* $e_t^g = 0$ *and* $e_t^g = \tau > 0$ *and let* $(y_t^*(\omega))_{t=0}^{\infty}$ *and* $(y_t'(\omega))_{t=0}^{\infty}$ *be the corresponding income distributions. If τ is not too large, then for each generation t, $t = 0, 1, \ldots$, the income distribution $y_t'(\omega)$ is more equal than the income distribution $y_t^*(\omega)$.*

An implication of Proposition 3 is that the introduction of a compulsory education results in a more equal intragenerational distribution of the human capital for all periods. We claim, without a proof, that the CE with compulsory education, τ, converges to a steady state. Denote the human capital distribution at this steady state by $h^\tau(\omega)$. The initial human capital distribution, $h^0(\omega)$, can be considered as the steady-state distribution with $\tau = 0$. The next result shows that for τ not 'too large' we can guarantee that the level tail of the distribution of $h^0(\omega)$ is shifted to the right when we introduce compulsory education with level τ.

Corollary. Assume that the initial steady-state distribution of human capital, $h^0(\omega)$, has a support $[m, M]$ [where $\inf h^0(\omega) = m$]. There exists $\tau^* > 0$ such that for any $0 < \tau < \tau^*$ the support of $h^\tau(\omega)$ is $[m + \varepsilon^*(\tau), \hat{M}(\tau)]$ where $\varepsilon^*(\tau) > 0$ and $\hat{M}(\tau) > M$.

Thus, compulsory schooling is a policy that improves the situation of the very poor fraction of the population. Social public policy tries to guarantee a minimum standard of living by using various intervention methods. The above corollary shows that compulsory schooling is an effective policy to achieve this goal. It is interesting to note that this policy, in addition, implies a transition to a steady state with a better distribution of income and a higher aggregate output for the economy.

5. Welfare implications

Since compulsory schooling, financed by tax on income, constitutes, basically, some transfer from individuals with higher human capital to

individuals with lower human capital, we cannot expect it to result in a Pareto improvement, at least not at the early stages. Consider a CE with taxes at rate τ financing compulsory schooling. We say that CS is *acceptable* be generation t if the majority of people in G_t prefer the equilibrium with CS to the one without it. The compulsory schooling is *acceptable in the long run* if there exists some $T < \infty$ such that CS is acceptable by all generations t for $t \geq T$. Thus, in the long run CS will prevail once it is established where in each period it is approved according to majority voting. The voters are only the young since the old generation is indifferent.

Proposition 4. Consider a competitive equilibrium with compulsory education e^g. If e^g is not too large, than this CE is acceptable in the long run. That is, for some finite T the majority of people in each generation, G_t, $t \geq T$, will prefer the compulsory schooling regime.

One should be careful in interpreting the proposition. Basically, it provides some comparison of the steady states (if they exist) with and without compulsory schooling. This result is not about the political support for public education because generations born after T are not asked to vote on something already implemented. The issue is whether people at date $t \geq T$ will vote for *maintaining* the compulsory schooling regime from t onwards. This condition need not be satisfied. Also it is not necessarily true that generation 0 will vote for compulsory schooling. Whether G_0 will vote for compulsory schooling (i.e. $T = 0$) depends on the shape of the initial distribution of human capital, $h_0(\omega)$. For example, if the mean of $h_0(\omega)$ is much larger than the median, it is very likely that the majority of individuals in G_0 are better off under the CS regime and hence vote for it at $t = 0$.

Compulsory schooling has two affects: the redistributional role and the welfare-enhancing role. Since we do not make any specific assumption regarding the skewness of the human capital distribution, the redistributive affect alone may not be supported by the majority of voters. However, since this measure corrects the externality at least partially, then from this aspect it is desirable by the majority. When the population is more homogeneous, the latter affect dominates and hence the majority will support this education scheme. Let us emphasize, however, that our statement in Proposition 4 is normative and is not a statement about the outcome of the political process.

6. Conclusions

The analysis presented here suggests that compulsory schooling, which is financed by proportional taxes on income, is a public policy that enhances growth, makes the distribution of earnings more equal while the majority of the population is better off in the long run under this regime. As a result, the

wide implementation of this policy around the world can be explained by its role in achieving a preferred allocation of resources and more equal distribution of human capital.

We show in our framework that a higher growth path for the economy is accompanied by a more equal distribution of income.[16] This result is obtained by comparing two different paths of growth and it is compatible to evidence of a cross-section between countries. The question whether along the growth path income becomes more equal cannot be answered when income distribution is endogenous. The reason is that whenever the income distribution approaches some steady state with positive frequency on several income levels, the change in the distribution may depend on the initial conditions which are given exogenously. On the other hand, a model where the income distribution approaches full equality cannot be considered as a model that endogenously determines the income distribution. Hence, the question of variations in income inequality along a growth path should be studied by using a comparative analysis of equilibrium growth paths and the associated income distributions as done in this paper.

Finally, we have shown that the compulsory education regime is 'supported' by the majority of people in G_t only for $t \geq T$, where $T = T(e^g)$, $T(e^g) \to 1$ as $e^g \to 0$. The conditions which would provide a positive theory for the implementation of CS are still an open question in our model.

Appendix

Proof of Proposition 2. Let us consider first the case where $e_t^g = 0$ for all t, i.e. $\tau = 0$. From (24), since $\bar{h}_t \equiv h_t$ and $\bar{h}_t' = 0$ in this case we obtain that

$$\frac{\alpha_3}{\alpha_4(\omega)} \frac{A(e_t^*)}{A'(e_t^*)} \geqq 1 - e_t^*(\omega) \text{ with equality if } e_t^*(\omega) > 0. \tag{33}$$

Consider an individual who chooses $e_t^*(\omega) = 0$ when $\tau = 0$. For this individual the change to $\tau > 0$ has no effect on e_t; that is, the optimum remains $e_t(\omega) = 0$. Rewriting (24) for all ω, where $e^g + e_t(\omega) > 0$, we obtain

$$\frac{A(e^g + e_t(\omega))}{A'(e^g + e_t(\omega))} = \frac{\alpha_4(\omega)}{\alpha_3} \frac{1 - e_t(\omega)}{1 - \beta(1 - e_t) \dfrac{e^g(h_t(\omega) - H_t^g)}{e^g H_t^g + e_t h_t}} . \tag{34}$$

In what follows we shall denote by a prime the CE when $\tau = e^g$ is positive and by an esterisk the CE when $\tau = 0$. Since $A(\cdot)/A'(\cdot)$ is an increasing

---

[16]Persson and Tabellini (1991) provide another model that has similar results. They also provide cross-country evidence that supports the implication that equality is positively associated with growth.

function and the RHS of (34) is an increasing function in $(1 - e_t)$, it is easy to verify (by way of negation) that if τ is not too large:

$$e^g + e'_t(\omega) > e^*_t(\omega), \quad \text{for all } \omega. \tag{35}$$

Given K_0 and $h_0(\omega)$, let us assume that when $e^g_t > 0$ (but not too large). Since $w'_0 = F_2(K_0, (1 - \tau)L_0) > w_0$ hence $K'_1 > (1 - \tau)K^*_1$ since by (33) $s'_0(\omega) > (1 - \tau)s^*_0(\omega)$ for all ω. For some $\mu_1 > 0$ we can write

$$L'_1 = (1 - \tau)\int A(e'_0(\omega) + e^g)h_0(\omega)^\beta \geqq (1 - \tau)(1 + \mu_1)\int A(e^*_t(\omega))h^\beta_0(\omega)$$

$$= L^*_1(1 - \mu_1)(1 - \tau).$$

Thus $L'_1 \geqq (1 - \tau)(1 + \mu_1)L^*_1$. Thus for some $\lambda^* > 0$:

$$F(K'_1, L'_1) \geqq (1 - \tau)(1 + \lambda^*)F(K^*_1, L^*_1).$$

Claim. For some $\hat{\varepsilon} > 0$ (which depends upon τ), for all $t \geqq 1$,

$$\int A(e'_t(\omega) + \tau) \, d\mu_t \geqq (1 + \hat{\varepsilon}) \int A(e^*_t(\omega)) \, d\mu_t. \tag{36}$$

Proof of the claim. By (33) and (34) we derive that for all ω, where $e^*_t(\omega) > 0$ and $e'_t(\omega) > 0$, we have

$$\frac{A(e^*_t)}{e^*_t A'(e^*_t)} = \frac{\alpha_4}{\alpha_3}(1 - e^*_t)/e^*_t,$$

$$\frac{A(e'_t + e^g)}{(e'_t + e^g)A'(e'_t + e^g)} \geqq \frac{\alpha_4}{\alpha_3} \frac{1 - e'_t + \hat{\lambda}_\tau(1 - e'_t)}{e'_t + e^g}.$$

Thus, using our assumptions about $A(\cdot)$ we conclude that

$$\left(\frac{\alpha_4(\omega)}{\alpha_3} + \hat{\lambda}_\tau\right)\frac{1 - e'_t(\omega)}{e'_t + e^g} \leqq \frac{\alpha_4(\omega)}{\alpha_3}\frac{[1 - e^*_t(\omega)]}{e^*_t(\omega)}, \quad \text{for all } \omega.$$

Hence, noting that $e^g \equiv \tau$,

$$\left(\frac{\alpha_4(\omega)}{\alpha_3} + \hat{\lambda}_\tau\right)\frac{1 + \tau}{e'_t + \tau} - \hat{\lambda}_\tau \leqq \frac{\alpha_4(\omega)}{\alpha_3}\frac{1}{e^*_t(\omega)}.$$

But $1 + \tau > e'_t + \tau$ implies that whenever $e^*_t(\omega) > 0$ and $e'_t(\omega) > 0$,

$$e'_t(\omega) + \tau \geqq e^*_t(\omega)(1 + \tau), \tag{37}$$

which by integration proves the claim.

It can also be shown that for some $\xi_j > 0$, $j = 1, 2, 3, \ldots$, we have

$$\int h'_t(\omega) \, d\mu_{t-1} \geqq \prod_{j=1}^{t} (1 + \xi_j) \int h^*_t(\omega) \, d\mu_{t-1}. \tag{38}$$

Moreover, $\{\xi_j\}$ does not converge to 0 for a given $\tau > 0$. Thus for n large,

$$(1 - \tau)(1 + \hat{\varepsilon}) \prod_{j=1}^{n} (1 + \xi_j) > 1. \tag{39}$$

To complete the proof of the proposition let us note the following facts. (a) By (33) we see that $e^*_t(\omega)$ is determined (whenever it is positive) regardless of $h^*_t(\omega)$, as long as $\tau = 0$. (b) By (34) it is easy to verify, using the monotonicity of $A(x)/A'(x)$, that when $\tau_t > 0$ (given H^g_t) the optimal $e'_t(\omega)$ increases as $h'_t(\omega)$ increases; thus $\mathrm{Cov}(e'_t, h'_t) > 0$. Therefore we can derive the following inequalities:

$$L'_t = (1 - \tau) \int A(e'_t(\omega) + e^g)[\bar{h}'_t(\omega)]^\beta \, d\mu_t$$

$$\geqq (1 - \tau) \int A(e'_t(\omega) + e^g) \int [\bar{h}'_t(\omega)]^\beta \, d\mu_t. \tag{40}$$

To get the last inequality we use the fact that $(\bar{h}'_t)^\beta \geqq \frac{1}{2}(h^*_t)^\beta$, to be proved during the proof of Proposition 3.

By (37) and (38) it can be verified that the RHS of (40) becomes larger than L^*_t for t large enough. From (23) we derive that: $K'_{t+1} = [\alpha_2/(\alpha_1 + \alpha_2)]w'_t L'_t$, $t = 0, 1, 2, \ldots$. By our assumption about the elasticity of F_L we have $w'_t L'_t > w_t L^*_t$, since $L'_t > L^*_t$. This proves our assertion about the capital stocks. $\square$

Proof of Proposition 3. Since there are no intergenerational transfers of capital,

$$y^*_t(\omega) = w^*_t h^*_t(\omega), \quad t = 0, 1, 2, \ldots, \tag{41}$$

$$y'_t(\omega) = (1 - \tau) w'_t h'_t(\omega), \quad t = 0, 1, 2, \ldots. \tag{42}$$

Let us prove the theorem by induction on t. At $t = 1$ we have $h^*_0(\omega) = h'_0(\omega)$ for all ω. Also by the concavity and strict monotonicity of $A(\cdot)$, using (35) we find that $A(e'_0(\omega) + \tau)[h_0(\omega)]^\beta$ is more equal than $A(e^*_0(\omega))[h_0(\omega)]^\beta$ [because it dominates it in the second degree stochastic dominance; see Rothschild and Stiglitz (1973)]. Therefore $y'_1(\omega)$ is more equal than $y^*_1(\omega)$. This clearly implies that $h'_1(\omega)$ is more equally distributed than $h^*_1(\omega)$. To continue this induction let us prove:

Lemma 1. Let $X(\omega)$, $Y(\omega)$, $\bar{X}(\omega)$ and $\bar{Y}(\omega)$ be positive random variables with c.d.f.s F, G, $\bar{F}$ and $\bar{G}$ correspondingly, each having a support $[a,b]$, $0 < a < b < \infty$. Define $Z(\omega) = X(\omega)Y(\omega)$ and $\bar{Z}(\omega) = \bar{X}(\omega)\bar{Y}(\omega)$. If $X \geq_1 \bar{X}$, $\|Y\| = \|\bar{Y}\| = 1$ and $Y >_2 \bar{Y}$, then $Z >_2 \bar{Z}$.

Proof of Lemma 1. Since the support of all the given four random variables is $[a,b]$ let us compute the c.d.f. of Z as follows:

$$H(\xi) = \text{Prob}\{Z(\omega) \leq \xi\} = \text{Prob}\{Y(\omega) = \theta \quad \text{and} \quad X(\omega) \leq \xi/\theta, \quad \text{for}$$
some $a \leq \theta \leq b\}$.

Hence we can write (assume $b/a \geq b$)

$$H(\xi) = \int_a^{b/a} G'(x)F\left(\frac{\xi}{x}\right)dx = \int_a^b G'(x)F\left(\frac{\xi}{x}\right)dx.$$

We shall use the Rothschild–Stiglitz (1970) criteria for SDSD to prove our assertion. Let $\bar{H}(\xi)$ be the c.d.f. of $\bar{Z}(\omega)$ and assume that the support of H and $\bar{H}$ is $[a', b']$:

$$\Delta(t) = \int_{a'}^t [H(\xi) - \bar{H}(\xi)]\,d\xi = \int_{a'}^t \int_a^b \left[G'(x)F\left(\frac{\xi}{x}\right) - \bar{G}'(x)\bar{F}\left(\frac{\xi}{x}\right)\right] dx\,d\xi.$$

Since $F >_1 \bar{F}$ we have $F(\theta) \leq \bar{F}(\theta)$ for all θ and thus we can write

$$\Delta(t) \leq \int_a^b [G'(x) - \bar{G}'(x)]\left[\int_{a'}^t \bar{F}\left(\frac{\xi}{x}\right)d\xi\right]dx. \tag{43}$$

However, $\bar{F}(\xi/x)$ is positive and decreasing in x on $(a,b]$. Hence the function $m(x) = \int_{a'}^t \bar{F}(\xi/x)\,d\xi$ has the same properties as a function of x for all $a < t < b$. Using integration by parts, noting that $G(a) - \bar{G}(a) = 0$ and $G(b) - \bar{G}(b) = 0$, we obtain from (43) that

$$\Delta(t) \leq -\int_a^b [G(x) - \bar{G}(x)]m'(x)\,dx.$$

However, since $G >_1 \bar{G}$ we have $G(x) - \bar{G}(x) \leq 0$ and since $m'(x) \leq 0$ we have shown that $\Delta(t) \leq 0$ for all $t \in (a,b)$. This implies that $Z >_2 \bar{Z}$ [see Theorem 2.3 in Brumelle and Vickson (1975)], which completes the proof of Lemma 1.

To complete the proof of the theorem assume that for a given t the income distribution $y_t'(\omega)$ is more equal then $y_t^*(\omega)$. We shall use here Theorem 1 of Rothschild and Stiglitz (1973, p. 191). Thus to prove the induction step

notice first that by (41) and (42) our assumption implies that $h'_t \geqq_2 h^*_t$. Since $\bar{h}'_t$ is attained from h'_t by averaging it with its own average, $\bar{H}'_t$ (this is similar to a mean-preserving squeeze); hence, $\bar{h}'_t \geqq_2 h^*_t$. Since $0 < \beta \leqq 1$ we also obtain that $(\bar{h}'_t)^\beta \geqq_2 (h^*_t)^\beta$.

Using (35), $\int (e'_t(\omega) + \tau) > \int e^*_t(\omega)$ and A is concave. Hence $A(e'_t(\omega) + \tau)/\lambda'$ dominates SDSD $A(e^*_t(\omega))/\lambda$, where $\lambda' = EA(e'_t(\omega) + \tau)$ and $\lambda = EA(e^*_t(\omega))$. In particular this implies first degree stochastic dominance. Using Lemma 1 we find that $h'_{t+1}(\omega) = A(e'_t(\omega) + \tau)[\bar{h}'_t(\omega)]^\beta$ is more equal than $h^*_{t+1}(\omega) = A(e^*_t(\omega))[h^*_t(\omega)]^\beta$. This clearly implies that $y'_{t+1}(\omega)$ is more equal than $y^*_{t+1}(\omega)$, thus proving the induction step. $\square$

Proof of the corollary. Without loss of generality assume that in the no-intervention steady state, i.e. $\tau = 0$, case, $\inf e^*(\omega) = 0$. Thus, m solves the equation $h = A(0)h^\alpha$, namely $m = (A(0))^{1/1-\alpha}$. Now consider the steady-state distribution $h^\tau(\omega)$ when compulsory schooling at level $0 < \tau < \tau^*$ is introduced. Since $e'_t(\omega) + \tau > e^*_t(\omega)$ for all ω and t this inequality should hold in the steady state as well [proved as in the proof of Proposition 2, Eq. (40)]. Thus $\mathrm{essinf}[e'_t(\omega) + \tau] = \bar{e} + \tau > 0$. The infimum of $h^\tau(\omega)$, to be denoted by h_m, must be the solution of the equation

$$h_m = A(\bar{e} + \tau)\left[\frac{\bar{e}h_m + \tau\bar{H}^\tau}{\bar{e} + \tau}\right]^\beta,$$

where $\bar{H}^\tau$ is the average of $h^\tau(\omega)$ (hence $h_m \leqq \bar{H}^\tau$). By our earlier results $\bar{H}^\tau > \int h^0(\omega)$, and since $A(\bar{e}) > A(0)$ let us prove now that this infimum of $h^\tau(\omega)$, to be denoted by h_m, is strictly larger than m. Let us write

$$\left[\frac{h_m}{A(\bar{e})}\right]^{1/\beta} - \frac{\bar{e}}{\bar{e} + \tau}h_m = \frac{\tau}{\bar{e} + \tau}\bar{H}^\tau \geqq \frac{\tau}{\bar{e} + \tau}h_m.$$

Define $\bar{y}$ by the equation

$$\bar{y} = A(\bar{e})(\bar{y})^\beta.$$

However, by the above inequality h_m satisfies $h_m \geqq A(\bar{e})h_m^\beta$. Thus $h_m \geqq \bar{y}$. But $A(\bar{e}) > A(0)$ and hence $\bar{y} > m$, which proves our claim.

Proof of Proposition 4. We shall apply Propositions 2 and 3 to show that the majority of each G_t are better off with the CS equilibrium compared with the no-intervention case. It was proved in Proposition 2 that, given the CS at level $\tau = e^g$ (not to large), there exists some $N(\tau) < \infty$ such that for any date t, $t > N(\tau)$, the effective labor supply is higher, i.e. $L'^e_t \geqq L^*_t$; the total output in the CS equilibrium is higher, $F(K'_t, L'^e_t) > F(K^*_t, L^*_t)$ for all $t > N(\tau)$. Let us show first that the total income of each G_t, $t > N(\tau)$, is higher in the CS case.

By our assumptions about the production function as L_t^* increases to $L_t'^e$ (and K_t^* increases to K_t'), the aggregate income is $w_t' L_t'^e = L_t'^e F_L(K_t', L_t'^e) \geqq L_t^* F_L(K_t^*, L_t^*)$. Particularly, $\int y_t'(\omega) > \int y_t^*(\omega)$ for $t > N(\tau)$. Since the income distribution, $y_t'(\omega)$, is more equal than $y_t^*(\omega)$ for all $t \geqq 1$ for each α, $0 < \alpha < 1$, the percentage of the total income received by the lower-income 100α percent of G_t is higher in the CS equilibrium. However, for $t > N(\tau)$ the aggregate income of G_t is higher with the CS and hence the income of each individual in the lower-income 50 percent is higher under $y_t'(\omega)$ than under the distribution $y_t^*(\omega)$. As we have seen during the earlier proofs, the distribution of $h_{t+1}'(\omega)$ is more equal than the distribution of $h_{t+1}^*(\omega)$, and hence it is easy to verify that the lower-income 50 percent of the population in G_t are better off in the CS equilibrium for all $t > N(\tau)$. This means that the CS equilibrium is acceptable for all generations from $N(\tau)$ to ∞. $\square$

References

Aiyagari, S.R., 1989, Equilibrium existence in an overlapping generations model with altruistic preferences, Journal of Economic Theory 47, no. 1, 130–152.

Angrist, A.J. and A.B. Krueger, 1990, The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples, NBER Working paper no. 3521.

Angrist, A.J. and A.B. Krueger, 1991, Does compulsory school attendance affect schooling and earnings?, Quarterly Journal of Economics CVI.

Atkinson, A.B., 1970, On the measurement of inequality, Journal of Economic Theory 2, 244–263.

Azariadis, C. and A. Drazen, 1990, Threshold externalities in economic development, Quarterly Journal of Economics CV, 501–526.

Barro, R., 1974, Are government bonds net wealth, Journal of Political Economy 82, 1095–1117.

Becker, G.S., 1975, Human capital (University of Chicago Press, Chicago, IL).

Becker, G.S. and N. Tomes, 1979, An equilibrium theory of the distribution of income and intergenerational mobility, Journal of Political Economy 87, 1153–1189.

Bernheim, B.D. and K. Bagwell, 1988, Is everything neutral?, Journal of Political Economy 96, 308–338.

Brumelle, S.L. and R.G. Vickson, 1975, A unified approach to stochastic dominance, in: Ziemba and Vickson eds., Stochastic optimization models in finance (Academic Press, New York).

Chiswick, B.R., 1969, Minimum schooling legislation and the cross-sectional distribution of income, The Economic Journal 79, 495–507.

Dalton, H., 1920, Some aspects of the inequality of income in modern communities (George Routledge, London).

Diamond, P., 1965, National debt in a neoclassical growth model, American Economic Review 55, 1126–1150.

Fernandez, R. and R. Rogerson, 1991, On the political economy of educational subsidies, manuscript.

Friedman, M., 1962, Capitalism and freedom (University of Chicago Press, Chicago, IL).

Judd, K., 1985, The law of large numbers with a continuum of IID random variables, Journal of Economic Theory 35, 19–25.

Loury, G., 1981, Intergenerational transfers and the distribution of earnings, Econometrica 49, 843–867.

Lucas, R., 1988, On the mechanics of economic development, Journal of Monetary Economics 11, 3–42.

Melton, J.H., 1988, Absolutism and the eighteenth century origins of compulsory schooling in Prussia and Austria (Cambridge University Press, Cambridge).

Organization for Economic Cooperations Development (OECD), 1983, Compulsory schooling in a changing world (OECD, Paris).

Perotti, R., 1991, Income distribution and growth: Theory and evidence, Working paper, Columbia University.

Persson, T. and G. Tabellini, 1991, Is inequality harmful for growth?, manuscript.

Romer, P., 1986, Increasing returns and long run growth, Journal of Political Economy 19, 1002–1037.

Rothschild, M. and J. Stiglitz, 1973, Some further results on measurement of inequality, Journal of Economic Theory 6, 188–204.

Saint-Paul, G. and T. Verdier, 1991, Education, democracy and growth, manuscript.

West, E.G., 1970, Education and the state: A study in political economy (The Institute of Economic Affairs, London).