Economica (1999) **66**, 317–33

Educational Mobility: The Effect on Efficiency and Distribution

By Jan Eeckhout

The London School of Economics

Final version received 22 December 1998.

Within a microeconomic framework, educational mobility and inequality are studied. The labour market is characterized by imperfectly substitutable skills and production occurs in monopolistically competitive industries that exhibit local non-convexities. Education allows for upward mobility. It is shown that multiple mobility equilibria exist in the stage game. In addition, for some skill levels, Pareto improvements are possible through adjustment policies. In the repeated game, a sufficient condition is derived for polarization, in which case the economy exhibits a low growth path. A higher growth path can be achieved through intertemporal redistribution. Without adjustment, inequality will increase continuously.

Introduction

In this paper I investigate the effect of labour market mobility through the acquisition of education. I provide a detailed microeconomic structure of the economy, which relies on the existence of a local non-convexity in the production technology. In the presence of educational mobility, this non-convex production technology implies positive externalities as well as strategic complementarities. The main results that follow from this analysis are threefold. First, multiple Nash equilibria are shown to exist in the stage game, giving rise to a coordinating role for government. Second, with the positive externality, higher skilled workers benefit from mobility from below; this justifies the standard Pigouvian tax which is both efficiency- and equity-improving. Finally, in the dynamic context of the repeated stage game, it is shown that an industry's balanced growth rate is a function of the size of the industry. If higher-skill-level industries are larger, persistent polarization occurs for a large initial skill differential. If intervention does not bridge the gap now, it will be even more difficult in the future.

This is derived in the context of markets that have two main features. First, the degree of substitutability differs between consumption and production markets. Consumption goods of different quality are perfectly substitutable, whereas the input of skills in the production of quality is not substitutable at all. In some sense, this differential degree of substitutability is very much as in frictionless matching models (see the assignment game, Gale and Shapley 1962, or more recently Kremer and Maskin 1995), where the output (money) is perfectly substitutable whereas inputs (the skills associated with a worker) are not. Fernandez and Gali (1997) introduce education into such matching models. Second, the local non-convexity in production is modelled in a competitive economy with differentiated goods. Each good's industry exhibits increasing returns to scale in production. With free entry, the whole industry is monopolistically competitive while at the same time globally exhibiting returns to scale. Though skills are perfectly insubstitutable, an education technology allows for a change in the skill level. In the presence of the non-convexity, there will be positive externalities and strategic complementarities from educational mobility. The reliance on increasing returns in production and the resulting externalities and multiplicity of equilibria have widely been applied in economic geography (Krugman and Venables 1992). Recent work by Cooper and Corbae (1997) relies on similar non-convexities in order to study coordination problems in financial markets. The approach taken here tries to start from similarly realistic premises, i.e. monopolistic competition with non-convexities, in order to provide an endogenously derived micro structure for both externalities and strategic complementarities. A rigorous treatment of the necessary and sufficient conditions for externalities and complementarities is provided in Cooper and John (1988). Relying on the results of the matching, the economic geography and the coordination failure literature, our model provides a detailed micro structure within which inequality will be studied.

The study of persisting inequality has received quite a lot of attention. Most of the literature focuses on the relation between heterogeneous initial wealth levels and the evolving wealth distribution. In the presence of capital market imperfections, initial wealth levels can impose a constraint on the level of optimal investment. Several authors (e.g. Banerjee and Newman 1993, Galor and Zeira 1993, Piketty 1992 and Ferreira 1995) show that an ergodic distribution exits with a common technology. Here the premise is heterogeneity in ability or skill levels, whereas capital markets are assumed to be perfect. Not only the matching literature has considered heterogeneity in skills. Arnott and Rowse (1987), for example, study the effect of mixing heterogeneous students in classes. In the endogenous growth literature, Eeckhout and Jovanovic (1998) and Eicher (1996) are even closer to the approach followed here. They look at the interaction between the production technology (with externalities) and the incentives to invest in education. Externalities arise from production, whereas a constant-returns-to-scale education technology triggers this externality.

In the next section, the economy and its population are described in detail and the general equilibrium outcome without mobility is derived. Modelling this micro behaviour in detail involves some inevitable algebra. Section II discusses the impact of mobility in the stage game. It is shown that multiple equilibria can exist for a certain range of skill differentials. Apart from a coordination issue, there is also a problem of free riding, which can result in Pareto improvements by subsidizing education of the lower skilled people. Without subsdidy, the cost of investment is higher than the private return, so that no investment occurs. However, the high types benefit from their mobility, so they will be better off if they can induce low types to invest. In Section III the game is played repeatedly. This allows us to consider the effect of growth. Although this paper does not pretend to provide a theory of growth, with endogenous accumulation of human capital, the growth rate will depend on the distribution of skills and thus on education. The relation between equity and efficiency will be identified. In the tradition of the literature on ergodic distributions in the presence of capital market imperfections, the limiting distribution of both wealth and ability can be related to the growth of the economy. A sufficient condition is derived for which there is polarization, i.e. ever-increasing inequality. Some concluding remarks are made in Section IV.

I. The Basic Model

The economy is populated with heterogeneous individuals characterized by a type q. There is a finite number of n types indexed by q_i, $i = 1, \ldots, n$. Initially, they are distributed according to density $\phi(q_i)$, and the size of the whole population is normalized to one: $\sum \phi_i = 1$. Individuals are both producers and consumers and the type of an agent is determined by her productive ability or level of skill as a worker—the higher q_i, the more productive.

Workers produce in order to derive utility from consumption. To formalize the notion of less than perfect substitutability of labour, it will be assumed that workers of a certain type will only work in the same industry as workers of the same type. In fact, this implies complete insubstitutability of inputs in production. Some degree of substitutability is endogenously derived in models of perfect matching (e.g. Kremer and Maskin 1995) with complementarities between inputs. Here, similar characteristics (i.e. complete insubstitutability) could be derived endogenously in a matching model with a degree of complementarity (i.e. the cross-partial derivative between inputs) equal to infinity. Different degrees of substitutability would not alter any of the results below, provided that the degree of substitutability is lower in production than in consumption. With this simplifying assumption, an industry is then defined by the type of its characteristic worker q_i with production in each of the n parallel industries. In terms of the quantity produced in the different industries, technology is identical for industries of the same size. However, workers with higher skill levels will produce higher-quality goods. Although the quantity produced in two industries may be identical, the value will differ depending on the level of skills of its workers.

With technology of the quantity produced being identical for all industries, we can specify the technology of a generic industry given its size ϕ. As mentioned in the Introduction, a monopolistically competitive technology is assumed which will incorporate spillover effects at the production level. The production process is modelled following Markusen (1989), allowing for free entry and a zero profit condition. Each industry is characterized by a sector of diversified input goods and an output sector. A worker in a certain industry can work either in the input or the output sector.

The input sector exhibits increasing returns and the output sector, constant returns. Hence the industry as a whole has an increasing returns technology. In any industry, $L \in [0, \phi]$ of the workers will work in the output sector. The other $\phi - L$ will work in the increasing-returns-to-scale input sector, which consists of a number of equally sized firms, producing some variety r of the input good, in a monopolistically competitive environment. The more varieties, the greater the quantity of the composite input X good which is used in production of the output. This technology is as in Dixit and Stiglitz (1977):

$$(1) \qquad X = \left(\sum_r x_r^{\frac{\sigma - 1}{\sigma}} \right)^{\frac{\sigma}{\sigma - 1}},$$

where $\sigma(>1)$ is the (constant) elasticity of substitution between the different variety inputs. Every x_r, the quantity of variety r of the input, is produced with increasing returns to scale: $s_r = \alpha + \beta x_r$, where s_r is the amount of labour used in that firm. It now follows that the amount of the composite input produced is convex in the number of workers in the input sector, $\phi - L$.

The output sector exhibits constant returns, the technology of which is $Y = L^{\theta} X^{1-\theta}$, where Y is the quantity produced. It follows that the output Y of the industry as a whole is convex in the number of workers ϕ.

We solve the problem for a given industry of size ϕ:

$$\frac{X}{L} = \frac{1-\theta}{\theta} \frac{\omega}{p_X}, \tag{2}$$

$$L^{\theta} X^{1-\theta} - \omega L - p_X X = 0, \tag{3}$$

$$p_r = \frac{\sigma}{\sigma - 1} \beta \omega, \tag{4}$$

$$\frac{\alpha(\sigma - 1)}{\beta} = x_r. \tag{5}$$

Equation (2) is the profit maximization condition in the output sector, with p_X the price of the composite input and ω the wage (in terms of units of production) in the industry. In the presence of free entry, the zero profit condition is given by (3). In the input sector, each of the diversified firms r maximizes profits (equation (4)) with (5) the zero profit condition. This allows us to calculate the number of firms m in the sector, each producing the same amount x_r:

$$m = \frac{\phi - L}{\alpha \sigma}. \tag{6}$$

Because of the increasing returns in the production of the composite input, and typically for this Dixit–Stiglitz type of technology, the price p_r for $m > 1$ will be higher than the price index faced by the output producers:[1]

$$\begin{aligned} p_X &= m^{1/(1-\sigma)} p_r \\ &= \left(\frac{\phi - L}{\alpha \sigma} \right)^{1/(1-\sigma)} \frac{\sigma}{\sigma - 1} \beta \omega \end{aligned} \tag{7}$$

after substitution for (4) and (6). Finally, market-clearing implies:

$$\omega \phi = L^{\theta} X^{1-\theta}. \tag{8}$$

Equations (2), (3), (7) and (8) are independent and contain four unknowns: ω, p_X, L and X.

The preferences of individuals over the goods produced in each individual industry depend on the quality of the good produced and hence on the level of skills of the workers. The quality of a good will simply be indexed by q. As a result, n types of good will be produced with quality q_i, with $q_1 < q_2 < \cdots < q_n$. Agents all have identical preferences independent of their type. Utility is increasing in both quality q and the quantity consumed Y_q. They perfectly substitute quantity for quality: $U = \sum_q q Y_q$. This specification is as in Rosen (1981). Individuals are risk-neutral.[2]

In general, consumers will equate the marginal rate of substitution, q_i/q_j, to the price ratio of different quality goods. If good $q_i = 1$ is taken as the numeraire good, prices satisfy

$$(9) \qquad p_q = q.$$

From (9), we know that the value of the output for industries with the same number of workers will be in exact proportion to the relative quality levels. As a result, the value of the output of an industry q can be written as qY_q. In addition, the wages between different industries can be compared. Each worker in an industry receives a wage equal to the value of the quantity $\omega(\phi)$ of that industry's goods. Hence, with different values of goods (equation (9)), the monetary wages will be given by

$$(10) \qquad w_q(\phi) = q\omega(\phi).$$

The advantage of the specified preferences and the production technology is that, although wages depend on the size of the industry, the quality or skill impact is separable.

Within this framework, the actions of an individual are the choice of the level of education. Costly education will enable an individual to increase her level of human capital, given an initial endowment of skills. This is beneficial, since it allows her to produce a higher-quality good and hence to receive a higher price for her labour. The cost of education is increasing in the level at which skills are augmented and depends on the initial skill level q. In general, the cost of education is given by the separable cost function $F(\Delta, q) = f(\Delta)q$, with

$$(11) \qquad \Delta = \frac{\Delta q_i}{q_i}.$$

To ensure existence, the following restrictions are imposed. $f(\Delta)$ is strictly convex ($f_{\Delta\Delta} > 0$), $f_\Delta \geq 0$, $f_\Delta(0) = 0$, $f_\Delta(\infty) = \infty$ and $f(0) = 0$. Investment in general depends on the initial level of skills, but the returns to skill are assumed constant. Individuals will choose the level of education Δ in order to maximize utility, taking the strategy of all other players as given. A Nash equilibrium will then be a rule such that each individual chooses an optimal strategy, taking into account the optimal strategy of all other players.

II. The Results of the Stage Game

With the production sector specified above, it is crucial how the wage is affected by the size of the workforce in the industry. With a monopolistically competitive production industry which displays increasing returns to scale, there will certainly be a positive effect of ϕ on the quantity of output produced. Simultaneously, from the fact that there is free entry, profits are driven to zero and the entire output accrues to the workers in this general competitive equilibrium context. Consequently, not only is the quantity produced increasing in the size of the industry, but the value of the wage is increasing. This is shown in the following lemma.

Lemma 1. The wage in any industry is increasing in the size of the industry ϕ:

$$(12) \qquad \frac{\partial w}{\partial \phi} > 0.$$

Proof. See Appendix.

This entirely captures the notion of external effects in the production technology. Although the education technology will not inherently exhibit any external effect, the returns to education are affected by the wage. Hence, since the wage depends on the degree of mobility (which is a synonym for education in this context), the decision of an individual to invest will depend on the level of investment of the others.

Consider the simplest possible case of a two-period game with two types, q_1 (low) and q_2 (high). In the first period of the game, all ϕ_1 (in the two-industry case, set $\phi_2 = \phi$, hence $\phi_1 = 1 - \phi$) types q_1 can choose between two strategies.

I Invest in education at cost $F(\Delta, q_i)$[3] in order to achieve a level of human capital q_2; in the second period, the q_1 type will be able to work in the q_2 quality industry and receive w_2.

NI Make no investment and work in the low-quality production industry for the second period.

Note that in (11) Δ is defined as the gap between the quality (or skill) levels of the two industries. Since the cost of investment function is increasing in its argument Δ, the return on investment must be a function of the gap Δ. The return on investment behaves as in Figure 1.[4] The return on investment function V_q can be defined as $V_q = q(1+\Delta)\omega(\cdot) - RF(\Delta, q)$.[5]

If a low type decides to play (I), the return on investment depends on the behaviour of the other low types. As the wage is an increasing function of the size of the industry (Lemma 1), the return on investment is increasing in the number of types q_1 investing. There is a positive externality from mobility. Now, two extreme situations can be considered. First, all the q_1 types decide to invest, so the size of the new q_2 industry is $\phi_1 + \phi_2 = 1$. The return on

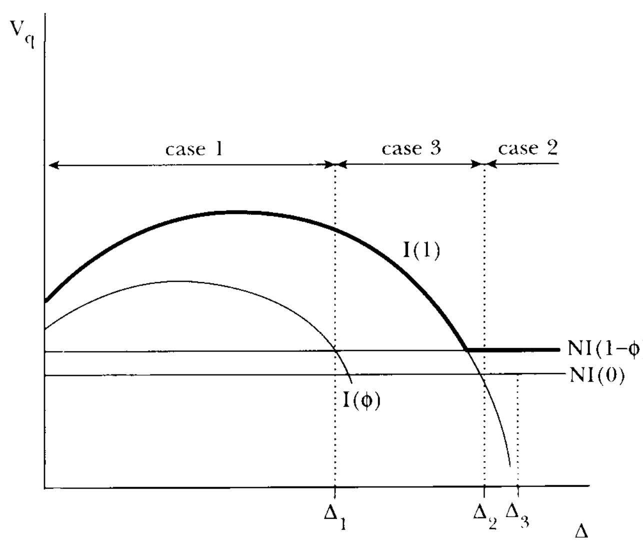

FIGURE 1

investment is

I(1): $q_2\,\omega(1) - RF(\Delta, q)$.
Second, none of the other low types decide to invest; the size of the q_2 industry does not change and the return on investment for a low type is
I(ϕ): $q_2\,\omega(\phi) - RF(\Delta, q)$,
which is always lower than in the first situation. Schedule I(1) is strictly above I(ϕ).

When a q_1 type decides not to invest, i.e. to play strategy NI, the return on investment function is independent of Δ because there is no cost. In the figure, this is represented by the horizontal lines NI. The wage received does depend on the strategy of the other workers, i.e. on how many types remain in industry q_1. Suppose no one invests: then the wage is equal to the current wage in industry 1, $w_1 = q_1\,\omega(1-\phi)$, given that there are $1-\phi$ workers. This is given by the schedule NI($1-\phi$). If all others invest and the decision is not to invest NI, then the wage is $w_1 = q_1\,\omega(0)$. This is represented by the schedule NI(0). NI(0) is below NI($1-\phi$). There is a negative externality from those who invest on those who decide not to invest.

The equilibrium strategies chosen by the q_1 types can now be analysed given an initial distribution, i.e. given $\{q_1, q_2\}$ and $\{\phi, 1-\phi\}$. Given the returns V_q for the strategies I(ϕ), I(1), NI($1-\phi$) and NI(0), and depending on the gap between the two types, there are three different cases.

Case 1: $q_2\,\omega(\phi) - RF(\Delta, q_1) > q_1\,\omega(1-\phi)$: *unique dominant strategy equilibrium.*
The equilibrium strategy for all $1-\phi$ types will be I. The motives are obvious: the disparity between the human capital levels—and hence the cost of investment *F*—is so low that they are better off when earning the high wage, irrespective of the strategies of the other q_1 types.

Case 2: $q_2\,\omega(1) - RF(\Delta, q_1) < q_1\,\omega(0)$: *unique dominant strategy equilibrium.*
The equilibrium strategy for all q_1 types will be NI. They do not invest, because the wage in the q_1 industry is higher than the wage in q_2 net of the cost of investment, even if all other low types would invest. Since the disparity between the different levels of human capital is so high, the cost of investment cannot be compensated by the gain in wage.

Case 3: $q_2\,\omega(\phi) - RF(\Delta, q_1) < q_1\,\omega(1-\phi)$ and $q_1\,\omega(0) < q_2\,\omega(1) - RF(\Delta, q_1)$: *multiple Nash equilibria.*
For Δ in this region there are: (i) a pure strategy equilibrium I for all types q_1; (ii) a pure strategy equilibrium NI for all types q_1; and (iii) a mixed strategy equilibrium[6] where all types q_1 are indifferent between I and NI. Each worker plays I with probability $\rho = \rho^*$, where

(13) $$\rho^* \in \{\rho \in (0,1): q_2\,\omega[\phi + \rho(1-\phi)] - RF(\Delta, q_1) = q_1\,\omega[(1-\rho)(1-\phi)].$$

The mixed strategy equilibrium is unstable because the slightest deviation leads to either one of the stable pure strategy equilibria (see also Figure 2). According to the Pareto criterion, I(1) always dominates NI($1-\phi$) and the mixed strategy equilibrium.[7]

Let us now consider some of the welfare implications. Case 3 illustrates that there is a serious problem of coordination failure. The equilibrium outcome where all workers choose I Pareto-dominates all other equilibria. However, the emergence of this equilibrium depends on the beliefs of all the other workers. There is a role for the government to improve coordination. Legislation that makes education mandatory up to a certain age can be interpreted as one such an example of coordinating action.

For the remainder of the paper, I will concentrate on the more interesting equilibria where no such coordination failure within an industry exists. Workers in different industries have different objectives. However, the interests within industries are identical for all workers. One way to think about this is that all workers are represented by a guild and decisions are made collectively. As a result, there is no coordination failure within groups of workers of the same type. In terms of Figure 1, this implies that only the highest curve of both I and NI is chosen. The equilibrium strategy when there is coordination within industries is then the upper envelope of I(1) and NI($1-\phi$).

After abstracting from the problems of coordination failure, one very substantial welfare issue remains. The game is designed such that only the lower human capital types choose a strategy. The higher types remain idle in the first stage. However, because of the externality arising from increasing returns, mobility of the low types will have an effect on the wage of the high types and thus on their utility. Closer inspection of the externality shows that, in fact, the low mobile types receive not the marginal product of their entry into the higher-quality industry, but the average product, which is lower than the marginal product. Since the higher types receive the average product as well, they benefit unequivocally from entry by the low types. There will be a case for a

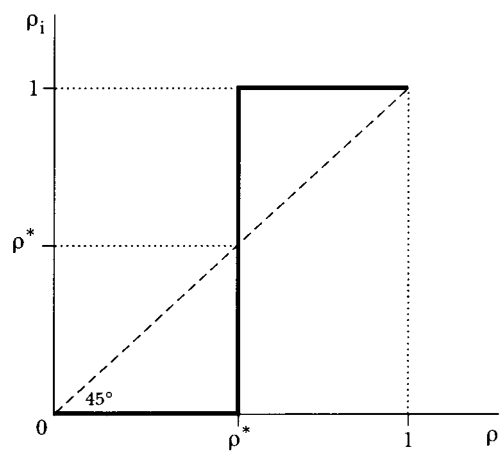

FIGURE 2

Pareto-improving subsidy. Two conditions have to be satisfied. (1) Case 2 must apply (i.e. $q_2\omega(1) - RF(\Delta, q_1) < q_1\omega(0)$, so that without subsidy NI is the equilibrium strategy. This condition is fulfilled for $\Delta > \Delta_2$. As a result, there will be no positive externality on the high types. (2) The subsidy must be large enough to induce the low types to make the investment: $q_2\omega(1) - RF(\Delta, q_1) + S(1-\phi)^{-1} = q_1\omega(0)$. This in fact implies that the high types give the minimum subsidy, S, necessary to achieve investment. There will obviously be an upper bound to the amount the high types are willing to subsidize. $S_{\max}$ is defined as the subsidy that makes the higher types indifferent between subsidizing education with the induced mobility (and hence with the resulting higher wage) and not subsidizing:

(14) $$S_{\max} = \phi(q_2\omega(1) - q_2\omega(\phi)).$$

In fact, $S_{\max}$ is the total value of the pure economic rent. It follows that only for $S \in [0, S_{\max}]$ are Pareto improvements possible. In terms of Δ, we can establish that Pareto improvements are possible for $\Delta \in [\Delta_2, \Delta_3]$, where Δ_3, provided it exists, is defined as

(15) $$\Delta_3 \in \{\Delta > \Delta_2 : q_2\omega(1) - RF(\Delta, q_1) + S_{\max}(1-\phi)^{-1} = q_1\omega(0)\}.$$

Since the external effect has an industry-wide impact, it is crucial that the effort of the high types to provide additional incentives to the low types is coordinated. There is indeed a serious free rider problem, which is not necessarily ruled out under the assumption of coordination within industry, but which cannot avoid deviation when there are incentives to do so. As a result, there will be a role for the government to impose a tax on the q_2 types and to subsidize education of the q_1 types if Δ is in the relevant interval.

III. The Repeated Game

The point of interest in this section is how the distribution evolves over time and how mobility can have an impact on both distribution and efficiency. We will study these aspects along the balanced growth path. Education will be the sole motor behind growth. It is therefore essential that now all types, including the highest types, have access to this investment in education technology.

Consider the two-period game from Section I (still with two types: $n = 2$), which is repeated in a successive-generations model where every parent gives birth to one child at the end of the second period. Monetary bequests are left out of the analysis, because capital markets are assumed to be perfect and as a result there will be no effect of the distribution of wealth on investment opportunities. Children inherit the human capital that the parent has accumulated in the second period of her generation. It follows that human capital is inheritable and accumulatable, and as a result the distribution of human capital at the end of one generation is reproduced at the beginning of the next generation.

In this framework with growth, both high types and low types will invest in education, irrespective of the mobility issue. Abstracting for the moment from the possibility of mobility, the problem for a type q will be to choose the amount of investment in education such that: $\Delta^* \in \arg\max\{V_q\} = \{w_{q,t+1}(\phi(q), \Delta) - RF(\Delta, q)\} = \{q(1+\Delta)\omega(\phi(q)) - RF(\Delta, q)\}$. But this clearly abstracts from the coordination problem within the industry. In that case there would be a continuum of Nash equilibria. We will consider the behaviour within an industry as that of a guild implying coordination between the workers in an industry. In general, the optimal amount of investment is given as

(16) $$\Delta^* = F_{\Delta}^{-1}\left(\frac{1}{R}\,\omega(\phi)q\right).$$

In order for the investment technology to be independent of the level of skills q, F is chosen to be multiplicatively separable ($F(\Delta, q) = f(\Delta)q$). The technology $V_q = q(1+\Delta)\omega(\phi) - Rf(\Delta)q$ can then be decomposed into a component v, independent of the initial human capital q, and a component q: $V_q = v(\Delta)q$. Hence, $v(\Delta) = (1+\Delta)\omega(\phi) - Rf(\Delta)$. Note that, with reference to the optimal taxation literature (Mirrlees 1971), the marginal rate of substitution is independent of type q and as a result there is no single crossing (i.e. the marginal rate of substitution is increasing in type). As a result, the optimal choice Δ^* is independent of q:

(17) $$\Delta^* = f_{\Delta}^{-1}\left(\frac{1}{R}\,\omega(\phi)\right).$$

The intuition behind this investment technology is as follows. As Δ is a measure for the gap between two levels of human capital—q after investment in education and q at the beginning of period 1, i.e. the initial level of human capital—Δ^* is the optimal amount by which to augment the level of skill q. At the same time, Δ (see equation (11)) is a measure for a percentage increase in the initial human capital, and hence the outcome after investment depends upon the initial level of human capital q. In other words, although the optimal amount of investment is independent of q (constant returns to skill), the level of human capital after investment is the initial-level q augmented with Δ^*.

For a given Δ, the cost of investment function $F(\Delta) = f(\Delta)q$ is increasing in q. This may at first sight seem to indicate that this technology exhibits decreasing returns to skills, which is entirely at odds with the empirical findings. However, as Δ is the ratio with denominator q and as f is convex, the cost of investment function F exhibits increasing returns to skill. Similar assumptions are made in work on endogenous growth models; see e.g. Eeckhout and Javonovic (1998).

The optimal investment decision is illustrated in Figure 3 for different-sized industries. Consider one curve, representing the return on investment of an industry for a given ϕ. Since the cost of investment function $f(\cdot)$ is strictly convex in Δ and the gains from investment $(1+\Delta)\omega(\cdot)$ are linear in Δ, the return on investment function $v(\Delta)$ is strictly concave, for $\Delta \in \Re^+$. As a result, there will be a unique solution for Δ^*. The solution always exists because $f_{\Delta}(0) = 0$, $f(0) = 0$ and (given the domain of f is $\Re^+$) infinitesimal amounts of investment have infinitely large returns which makes some investment always attractive. This is the mechanism which results in a strictly positive growth rate of an industry.

Given this technology, it follows that—still abstracting from the possibility of two groups merging and maintaining the absence of coordination problems within industries—the chosen amount of investment of an industry will be the

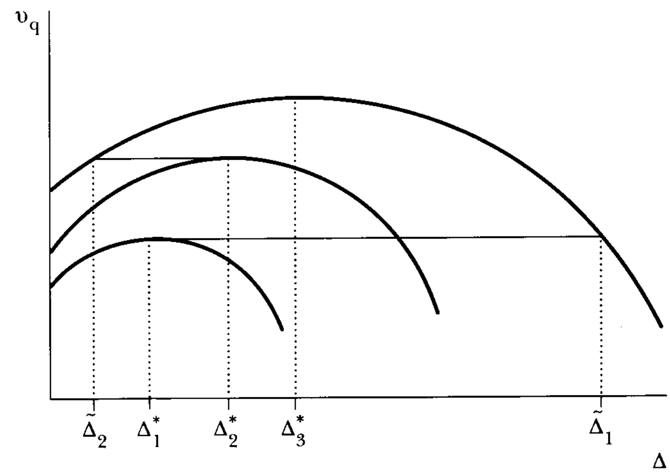

FIGURE 3

optimal amount Δ^*. We can now compare the growth rates of different industries, which by definition coincide with the amount of investment Δ. Because of the assumption of constant returns to skill, growth rates will, other things equal, be identical across different quality industries. However, in Proposition 1 it is shown that the growth rate depends on the size of the industry.

Proposition 1. The growth rate of an industry is increasing in the density of the workers with that level of human capital:

$$\frac{\partial \Delta^*}{\partial \phi} > 0.$$

Proof. See Appendix.

The intuition of Proposition 1 is shown in Figure 3. Since $\omega(\cdot)$ is strictly increasing, the return on investment function v is strictly higher for a higher density ϕ. Figure 3 gives v for three different densities, where $\phi_1 < \phi_2 < \phi_1 + \phi_2$. Proposition 1 is also illustrated graphically with $\Delta_1^* < \Delta_2^*$. From the proposition, it also follows that an industry will have a maximal growth rate when all workers in the economy work in the same industry, i.e. when $\phi = 1$. So far, we have concentrated on the optimal investment when no mobility between industries was possible. However, in line with the results from Section I, the low types may be willing to invest more than Δ^*, if they can join a higher-quality industry and thus benefit from the externalities from a larger workforce. As I have shown, the high types too benefit from the externality. Translated to the repeated game with investment by both types, this means that the high types may be willing to invest less than the optimal amount. However, the willingness to over/underinvest is bounded by the outside option, i.e. the return when industries do not merge. The maximum/minimum individuals are willing to invest has to make them at least as well off as in the case of no industries merging. Hence the following definition, which is also illustrated in Figure 3.

Definition 1.

$$\tilde{\Delta}_1 = \{\Delta \in \Re^+ : \omega(\phi_1)(1+\Delta_1^*) - Rf(\Delta_1^*) = \omega(\phi_1+\phi_2)(1+\Delta) - Rf(\Delta),\ \Delta > \Delta_1^*\}$$

(maximal investment by q_1 over and above Δ_1^* which makes the individual indifferent between the return in the merged (large) industry and the optimal return in the separate industry);

$$\tilde{\Delta}_2 = \max\{0, \Delta \in \Re^+ : \omega(\phi_2)(1+\Delta_2^*) - Rf(\Delta_2^*) = \omega(\phi_1+\phi_2)(1+\Delta) - Rf(\Delta),\ \Delta < \Delta_2^*\}$$

(minimal investment by q_2 below Δ_2^* which makes the individual indifferent between the return in the merged (large) industry and the optimal return in the separate industry).

We can now establish how the distribution will evolve over time and derive a sufficient condition for the limiting distribution.

Proposition 2. A two-industry economy with constant returns to skills will remain polarized into the two industries with a continuously decreasing ratio of human capital q_1/q_2, tending to zero at infinity, if both the following conditions hold:

(a) $\phi(q_1) < \phi(q_2)$ (necessary condition);
(b) $q_1(1+\tilde{\Delta}_1) < q_2(1+\tilde{\Delta}_2)$ (sufficient condition).

Proof. see Appendix.

The intuition behind Proposition 2 is as follows. If industry 2 is larger (condition (a)), it will grow faster than industry 1 (from Proposition 1). At the same time, the initial disparity between the two industries is so large that, even after the q_1 types have invested maximally and the q_2 types minimally, they are still not near enough to merge. In that case, the decentralized economy will not merge. The next period, the gap between the two industries is even bigger, because the larger industry grows faster because of condition (a). It follows that the gap increases over time.

Corollary 2.1. Given the conditions of Proposition 2. If the lower types do not merge with the higher types now, they will never do so.

This follows from the proof of the proposition. It is shown that condition 2 holds even more strongly in the next period; hence no mobility will occur. This applies for all consecutive periods, so mobility will never occur. Proposition 2 provides us with a sufficient condition for no mobility in a decentralized system. The result of the proposition can easily be generalized to an n-industry economy, as long as Proposition 2 holds between every one of the neighbouring industries. Hence we can formulate a general definition for a steady state.

Definition 2. An n-industry economy is in a steady state if either:

(a) Proposition 2 applies $n-1$ times, between every industry i and $i+1$, $i = 1, \ldots, n-1$, $n > 1$; or
(b) $n = 1$.

Once the steady state does occur, there will be no changes in the growth rates of the different industries. Since Proposition 2 requires the higher-quality industries to be larger, they will grow at a faster rate (from Proposition 1). This gives rise to Proposition 3, concerning the inequality of the evolving distribution.

Proposition 3. In the steady state of an n-industry economy ($n > 1$), the distribution of skills of the current generation Lorenz-dominates the distribution of next generation.

Proof. See Appendix.

This means that there is an unambiguous increase in inequality over time, in the sense that the Lorenz curves of any two consecutive generations do not intersect.

Corollary 3.1. There is an unambiguous increase in the inequality of income.

Net income is given by $V_q = q\nu(\Delta^*)$. From condition (a) in Proposition 2, ϕ has to be increasing in q in the steady state, and from Proposition 1, ν will be increasing in q. As a result, if the distribution of q becomes more unequal, the distribution of $q\nu(\Delta^*)$ will become even more unequal.

IV. Conclusion

In this paper, a detailed microstructure of the labour market is provided. The economy is characterized by heterogeneously skilled workers, whose inputs in production are imperfectly substitutable, and by a production technology that exhibits a local non-convexity. Within this framework, education is a technology that allows workers to upgrade their skill levels. The main insight is that, apart from leading to a higher level of consumption, educational mobility also gets the economy on to a higher growth path, where the growth rate is related to the degree of income inequality.

In the static game, coordination problems exist. Legislation on compulsory schooling can be considered as a way of coordinating strategies of agents. Even in the case of perfect coordination, however, there is no mobility for a certain range of skill in the decentralized system, even though Pareto improvements can be achieved. This is because the higher types experience a positive externality from mobility. There is a free rider problem which cannot just be solved by inducing coordination.

In order to derive these results, I have made strong assumptions about the degree of substitutability. Types of worker (inputs in production) are not substitutable at all, whereas output is perfectly substitutable. This has been crucial for deriving the results. Relaxing these assumptions for a simplified version of the model shows however that the results still hold as long as the degree of substitutability of inputs is lower than the degree of substitutability of outputs.

When the basic game is repeated and the education technology is embedded in a constant growth economy, it emerges that the larger industries will grow faster. A sufficient condition for polarization resulting from a lack of mobility is derived. As the growth rate of an industry is increasing in the size of the industry, this steady state—i.e. no mobility—exhibits a low growth path. Moreover, over time, inequality in the economy increases unequivocally. The underlying reason for increasing inequality is the spread in skills. Mobility is too costly and the emergence of a poverty trap is possible. A higher growth path—which increases intertemporal social welfare—can be achieved, but it requires an intertemporal redistribution from the high types in the current generation to all types in the future generations. Since the cost of adjustment is a function of the spread between the skill levels, and since inequality increases over time, adjustment becomes more costly the longer it is postponed.

Although it is dangerous to derive policy implications from a rigid model providing a simplified representation of reality, some general guidelines may be useful. In the presence of some externality in production, i.e. when a compact distribution is more efficient, the government has a role to encourage maximal mobility. A long-term concern is the fact that no action now may cause irretrievable damage later. This is a particularly difficult dilemma because it involves intertemporal redistribution between the generations which is extremely costly for the currently skilled workers and thus is an unpopular measure to impose.

APPENDIX

Proof of Lemma 1.

The system of equations can be simplified and yields an explicit solution for ω as a function of ϕ as follows. Define

(A1) $$A=\left(\frac{\phi-L}{\alpha\sigma}\right)^{\frac{1}{1-\sigma}}\frac{\sigma}{\sigma-1}\beta.$$

Using (7), we can rewrite (2) and (3) respectively as

(A2) $$X=\frac{1-\theta}{\theta}\frac{1}{A}L,$$

(A3) $$L^{\theta}X^{1-\theta}-\omega L-A\,\omega X=0.$$

Equations (A2) and (A3) then yield

(A4) $$\left(\frac{1-\theta}{\theta}\right)^{1-\theta}A^{\theta-1}-\frac{\omega}{\theta}=0.$$

Equations (8) and (A2) can be written as

A5) $$\omega\phi=L\left(\frac{1-\theta}{\theta}\right)^{1-\theta}A^{\theta-1}.$$

Equations (A4) and (A5) give us a very simple expression for L:

(A6) $$L=\phi\theta.$$

Using this and substituting in (A4)—i.e. rewriting the expression for A—we find the following explicit solution for ω as a function of ϕ:

(A7) $$\omega=\theta^{\theta}(1-\theta)^{1-\theta}\left(\frac{\phi(1-\theta)}{\alpha\sigma}\right)^{\frac{\theta-1}{1-\sigma}}\left(\frac{\sigma}{\sigma-1}\beta\right)^{\theta-1}.$$

Taking the partial derivative yields

(A8) $$\frac{\partial\omega}{\partial\phi}=\theta^{\theta}(1-\theta)^{1-\theta}\left(\frac{\sigma}{\sigma-1}\beta\right)^{\theta-1}\frac{\theta-1}{1-\sigma}\left(\frac{\phi(1-\theta)}{\alpha\sigma}\right)^{\frac{\theta-1}{1-\sigma}-1}\frac{(1-\theta)}{\alpha\sigma}.$$

Given α, β, $\phi > 0$; $0 < \theta < 1$; $\sigma > 1$; expression (A8) is positive. Since $w = \omega q$ and with $\partial q / \partial \phi = 0$, it follows that

$$\text{(A9)} \qquad \frac{\partial w}{\partial \phi} > 0. \qquad \square$$

Proof of Proposition 1.

Equation (17) is derived from

$$\Delta^* = f_{\Delta}^{-1}\left(\frac{1}{R}\,\omega(\phi)\right).$$

Derivation with respect to ϕ gives

$$\frac{\partial \Delta^*}{\partial \phi} = \frac{1}{f_{\Delta\Delta}} \frac{\partial \omega(\phi)}{\partial \phi} \frac{1}{R}.$$

From Lemma 1,

$$\frac{\partial \omega(\phi)}{\partial \phi} > 0,$$

and with $h(q)$ and $f_{\Delta\Delta}$ positive, it follows that

$$\frac{\partial \Delta^*}{\partial \phi} > 0. \qquad \square$$

Proof of Proposition 2.

It follows from Proposition 1 that condition (a) is necessary: the industry with the higher density has a higher growth rate. The condition is necessary because, if violated, industry 1 will grow at least as fast and thus there will be no decrease in the proportion of human capital. If the density at q_1 is strictly greater than at q_2, the proportion q_1/q_2 will increase and eventually the two industries will merge.

Condition (a) is not sufficient since there is the possibility that lower-type individuals will bridge the gap between their levels of human capital and start producing the q_2-quality good. Condition (b) refers to the case where the gap between the two levels of human capital is so high that no investment will be made to bridge the gap. In terms of the extended model of Section 1—i.e. the dynamic version with investment by all agents even without social mobility—Case 2 applies.

Since the cost of investment function is strictly convex in Δ and the gains from investment $\omega(\cdot)(1+\Delta)$ are linear in Δ, the investment function $V_q = vq$ is strictly concave for Δ in $\Re^+$. Moreover, $\omega(\cdot)$ is strictly increasing, so that Δ_1 and Δ_2 are uniquely defined. From condition (b), it follows that type q_1, when investing the maximal individually rational amount, will never be able to reach a level of human capital equal to that of type q_2, when the latter is investing the minimal individually rational amount.

Combination of the two conditions provides a sufficient condition for q_1/q_2 to be lower at the end of stage 2 compared with the beginning of stage 1. In the next generation, the distribution is exactly reproduced as it was at the end of stage 2, so that condition (a) remains unchanged and condition (b) will hold even more strongly because: (i) Proposition 1 implies that q_2 types will invest more than q_1 types, which will drive down the q_1/q_2 ratio; (ii) since Δ is independent of the level of human capital, $(1+\Delta_2)/(1+\Delta_1)$ will remain unchanged. This scenario will be repeated, and q_1/q_2 will continue to decrease over time. Over an infinite number of future generations, q_1/q_2 will tend to zero, since

$$\lim_{t \to \infty} \frac{q_1(1+\Delta_1^t)}{q_2(1+\Delta_2^t)} = 0. \qquad \square$$

Proof of Proposition 3

The underlying social welfare function according to which distributions are ranked is utilitarian. It will be shown that the Lorenz curve of this period's distribution is not lower than next period's for all q. Note further that the usual assumptions about the underlying welfare function and individual utilities apply. In terms of notation, next period's variables will be indicated by a prime.

Starting from observations

1. $\Delta_j^* > \Delta_i^*$, $\forall j > i$: by definition of steady state and from proposition 1,
2. $\phi(q_i) = \phi'(q_i)$: by definition of steady state, and
3. $q_i' = q_i(1 + \Delta_i^*)$,

we can show that the share of total income is never larger in the next period:

$$\text{(A10)} \quad \frac{\sum_1^k q_i \phi(q_i)}{\sum_1^n q_i \phi(q_i)} \geq \frac{\sum_1^k q_i' \phi'(q_i)}{\sum_1^n q_i' \phi'(q_i)}, \qquad \forall k = 1, \ldots, n.$$

Inverting (A10) and using observations 2 and 3 gives

$$1 + \frac{\sum_{k+1}^n q_i \phi(q_i)}{\sum_1^k q_i \phi(q_i)} \leq 1 + \frac{\sum_{k+1}^n q_i' \phi'(q_i)}{\sum_1^k q_i' \phi'(q_i)}, \qquad \forall k = 1, \ldots, n.$$

Dividing the numerator and denominator in the RHS through by $1 + \Delta_{k+1}^*$ gives

$$1 + \frac{\sum_{k+1}^n q_i \phi(q_i)}{\sum_1^k q_i \phi(q_i)} \leq 1 + \frac{\sum_{k+1}^n q_i((1 + \Delta_i^*)/(1 + \Delta_{k+1}^*))\phi'(q_i)}{\sum_1^k q_i((1 + \Delta_i^*)/(1 + \Delta_{k+1}^*))\phi'(q_i)}, \qquad \forall k = 1, \ldots, n.$$

From observation 1, it follows that $1 + \Delta_i^* > 1 + \Delta_{k+1}^*$, $\forall i > k + 1$, so that the numerator of the RHS is higher than the one on the LHS. Similarly, $1 + \Delta_i^* < 1 + \Delta_{k+1}^*$, $\forall i < k + 1$, resulting in the denominator on the RHS being smaller than the one on the LHS. Hence, the RHS is bigger than the LHS, which proves (A10). □

ACKNOWLEDGMENTS

I would like to thank Michele Boldrin, Marco Celentani, Raquel Fernandez, Francisco Ferreira, Gustavo Nombela, two anonymous referees and especially Kevin Roberts for discussion and comments.

NOTES

1. This transition mechanism through the price is the driving force of the Dixit–Stiglitz (1977) model. Since there is a 'taste' for variety in inputs, the nominal value p_r of an input is higher than the real value p_X. The greater the variety (i.e. the higher m), the lower the real value of the input.
2. Though there is no risk in this model, mixed strategy equilibria are investigated.
3. Capital markets are perfect, so investment can be paid for by borrowed money and will be repaid at rate $R = 1 +$ interest rate.
4. Figure 1 is a mapping of Δ onto the return on investment function for given densities $\{\phi_1, \phi_2\}$. (The figure is drawn for $\phi_1 < \phi_2$.) As Δ is defined as the relative gap between q_2 and q_1, the function indicates what happens in case of changing inequality.
5. Note that $q_1(1 + \Delta)\omega(\cdot) = q_2\omega(\cdot)$.
6. This mixed strategy equilibrium can analogously be interpreted as a pure strategy equilibrium where a fraction ρ decide to invest. At that point, a q_1 type is indifferent between I and NI.
7. Though I(1) dominates both the other equilibria, the equilibria NI$(1 - \phi)$ and the mixed strategy cannot be Pareto-ranked. In the mixed equilibrium, a fraction ρ will *ex post* be better off than in the all NI while a fraction $1 - \rho$ will be worse off.

REFERENCES

ARNOTT, R. and ROWSE, J. (1987). Peer group effects and educational attainment. *Journal of Public Economics*, **32,** 287–305.

BANERJEE, A. and NEWMAN, A. (1993). Occupational choice and the process of development. *Journal of Political Economy*, **101,** 274–98.

COOPER, R. and CORBAE, D. (1997). Financial fragility and the Great Depression. Mimeo, Boston University, June.

—— and John, A. (1988). Coordinating coordination failures in Keynesian models. *Quarterly Journal of Economics*, **103,** 441–64.

Dixit, A. and Stiglitz, J. (1977). Monopolistic competition and optimum product diversity. *American Economic Review*, **67,** 297–308.

Eeckhout, J. and Jovanovic, B. (1988). Inequality. Mimeo, NYU and UPF.

Eicher, T. (1996). Interaction between endogenous human capital and technological change. *Review of Economic Studies*, **63,** 127–40.

Fernandez, R. and Gali, J. (1997). To each according to. . . ? Markets, tournaments, and the matching problem with borrowing constraints. C.V. Starr Discussion Paper 97–11, New York University, February.

Ferreira, F. (1995). Roads to equality: wealth distribution dynamics with public–private capital complementarity. STICERD Discussion Paper, London School of Economics.

Gale, D. and Shapley, L. (1962). College admission and the stability of marriage. *American Mathematical Monthly*, **69,** 9–15.

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

Kremer, M. (1993). The O-ring theory of economic development. *Quarterly Journal of Economics*, **108,** 551–75.

—— and Maskin, E. (1995). Segregation by skill and the rise in inequality. Mimeo, Harvard University and MIT, March.

Krugman, P. and Venables, A. (1992). Integration, specialization and adjustment. CEP Discussion Paper no. 172, London School of Economics.

Markusen, J. (1989). Trade in producer services and in other specialized intermediate inputs. *American Economic Review*, **79,** 85–95.

Mirrlees, J. (1971). An exploration in the theory of optimal income taxation. *Review of Economic Studies*, **38,** 479–501.

Piketty, T. (1992). Imperfect capital markets and persistence of initial wealth inequalities. STICERD Discussion Paper no. 92/255, London School of Economics.

Rosen, S. (1981). The economics of superstars. *American Economic Review*, **71,** 845–58.