Discussion Paper Series

2003 – 08

Department of Economics
Royal Holloway College
University of London
Egham TW20 0EX

Stratification, Social Networks in the Labour Market, and Intergenerational Mobility

By Dan Anderberg* and Fredrik Andersson†

*Royal Holloway University of London, CEPR and CESifo; †Lund University

November 5, 2003

Abstract

This paper uses a model of human capital accumulation, labour market distortions, word-of-mouth communication, and community formation to analyse socio-economic stratification, educational choices and intergenerational social mobility. Workers obtain information about job opportunities from individuals in their local environment, implying that the social environment partly determines the expected returns to education. Stratified equilibria, when they exist, are characterised by low intergenerational social mobility and inefficient use of talent. The equilibrium responses to factors that generally encourage education may, in stratified outcomes, be highly asymmetric across socio-economic groups.

I Introduction

Recent decades have seen a pronounced increase in inequality in many countries, along with increasing wage returns to education. The increase in the returns to skill in many markets has encouraged participation in education. The response, however, has not been uniform across income groups. In the UK for example, the participation rate in higher education has more than doubled over the last twenty years; this general expansion has been accompanied by a strengthening of the education-parental income relationship, with participation in higher education rising much faster in the higher income groups than in the lower income groups (Blanden, Gregg and Machin, 2002). A similar development is manifest in the US; while there was a substantial increase in college attendance in the 1980s, the increase was much smaller among children from poor families than among children from rich families (Ellwood and Kane, 1999, and Acemoglu and Pischke, 2001). The uneven responses imply that further increases in inequality may be looming, and gaining an understanding of the causes of the asymmetric responses is of great importance if appropriate counteracting policies are to be designed.

The economic explanation traditionally put forward for the generally observed positive relationship between participation in education and parental income is the idea that poor families may be credit constrained (Loury, 1981 and Becker and Tomes, 1986). If children from poor families are unable to finance education, they will be in a weak position to respond to any increases in the returns to such investments. The evidence on the importance of credit constraints has been inconclusive so far, however; several different approaches and arguments have been put forward to get a handle on whether credit constraints affect educational choices.

Cameron and Heckman (1998, 2001) find that the estimated effects of family background on college attendance diminishes when scholastic ability is controlled for, suggesting that ability, rather than financial resources, determines college attendance.[1] This indicates that family background plays a role through its long-run effects on children's development including cognitive skills, not through short-term borrowing constraints. Carneiro and Heckman (2002) discuss the strength of other indirect evidence that has been put forward and conclude that this evidence is, in fact, often uninformative about credit constraints.

Insight into the question of the importance of credit constraints may also be obtained from the empirical literature on intergenerational mobility. Using US data, and exploiting information

[1]See also Mayer (1997) who uses several different approaches to get a handle on the effect of parental income on children's outcomes.

about expected financial transfers, Mulligan (1997) has recently argued that the evidence on intergenerational mobility does not lend strong support to the notion of significant borrowing constraints. A cross-country comparison of intergenerational mobility further reinforces the puzzle; from the sketchy evidence that is available it would appear that the correlation between fathers' and sons' long-run incomes is similar in the US and the UK, lying somewhere around 0.4 to 0.5 (Solon, 1999, and Dearden, Machin and Reed, 1997). Sweden, on the other hand, appears to have higher mobility than the US (Björklund and Jäntti, 1997),[2] while Checchi, Ichino and Rustichini (1999) have argued that Italy has lower mobility than the US. Since both the Italian and the Swedish educational system are primarily public, financed through taxation, thus reducing the privately borne cost of education, one would expect both countries to exhibit relatively high mobility. This suggests that there may be other important institutional features in an economy affecting intergenerational mobility.

A few studies have also tried to identify what changes, if any, in intergenerational mobility have accompanied the observed increase in inequality. For the US, Mayer and Lopoo (2001) and Fertig (2001) have found some evidence of increasing mobility. On the other hand, for the UK, Blanden et al. (2002) have found a striking decrease in intergenerational mobility in the recent decades, which they partly attribute to a strengthened connection between education and parental income.

The connection between parental income and children's economic success may, however, not only be due to financial constraints. It could e.g. be that children to rich parents grow up in advantageous social environments.[3] This insight has spawned a growing theoretical literature that looks at the causes and consequences of endogenous segregation when the social environment affects the human capital accumulation process. In this vein, the current paper shares Benabou's (1993) focus on the role of local and global interaction for stratification and educational choices. Benabou focuses on the simultaneous choice of education and residential choice, with local spillovers in education making those acquiring skills willing to pay to reside together. In contrast, in our framework youngsters choose education taking their locations (determined by the locational choices made by their parents) as given. In this respect our approach is more similar to Benabou (1996) who considers a model where parents, who differ in some character-

[2] See also the survey in Björklund and Jäntti (2000).

[3] See e.g. Blanden et al. (2002) for a discussion of different channels and the problems of empirically identifying direct and indirect channels.

istic, choose communities, and where a child's educational achievement depends on the parent's characteristic, on quality of the locally financed education, and on the composition of the community. Benabou examines various possible types of generic complementarities and shows how complementarities can amplify initial differences and make inequality more persistent across generations. Durlauf (1996) considers a model with endogenous sorting into communities where a child's educational achievement depends on locally financed education, the social environment, and on luck. Durlauf's main focus is on how endogenous segregation can generate persistent poverty. Our paper differs from the above mentioned papers in a couple of ways. First, while the above-mentioned papers assume perfect labour markets, in our analysis labour market imperfections and labour markets institutions is the key driving force generating spillover effects. Second, while in Benabou (1996) and Durlauf (1996) a key feedback mechanism from local social environment to individual choices is provided by local education finance (determined by some political mechanism). This is not the case in our model.

The current paper focuses on a particular role played by the social environment, namely its role in job search activities. A growing body of literature has documented the importance of social networks in finding employment (Granovetter, 1995 and Corcoran et al., 1989).[4] From a stratification and intergenerational mobility point of view, the role of social networks in job-finding processes is a potentially important, but hitherto fairly undeveloped, idea. From a cross-country point of view, the structure of the flow of information in labour markets can be expected to differ; e.g. public employment agencies appear to play a smaller role in the US than in many European countries.[5] Legislation also varies across countries; an extreme case is Sweden where employers are, in order to facilitate job-search, required to notify the National Labour-Market Board about any vacancies created, suggesting that comparatively few jobs may be allocated through social networks (Korpi, 2001).

In this paper we present a simple stylized model in which labour-market institutions – in particular, the mode of dissemination of information about scarce job-opportunities – constitute a driving force creating endogenous stratification and income correlation across generations. In the model, young individuals decide whether or not to acquire skills; however, due to imper-

[4] There is also a related emerging literature that looks at spatial structure and labour market outcomes. See e.g. Wasmer and Zenou (2001).

[5] Indeed, the research to date seems to have generated a widespread pessimism concerning the efficacy of U.S. public employment services (see e.g. Holzer, 1988) while the British findings give a more favourable picture of the public employment service (Gregg and Wadsworth, 1996).

fections in the labour market, skilled jobs are effectively being rationed. Information about job opportunities are spread by word-of-mouth communication, implying that an individual is more likely to hear about an opening in a skilled job if he has many skilled workers in his local environment. This generates a lower expected return to skill-acquisition for individuals from adverse social environments, making them less likely to invest in education. Based on the income obtained during working life, an individual then chooses a residential area, which defines the social environment of the individual's offspring. Using this setup we find that, while there always exists an equilibrium in which stratification does not occur, there may also exist equilibria with asymmetric or "stratified" communities. Equilibria with asymmetric communities are, however, never strictly superior from a social-optimality point of view; in addition, such equilibria will have lower intergenerational mobility than a non-stratified outcome.

The model allows us to consider the effects of various factors both on the existence of stratified equilibria and on behaviour starting from such equilibria. We find that reduced labour-market distortions and less reliance on word-of-mouth communication can stabilize a non-stratified equilibrium, while factors that generally encourage participation in education can be inherently destabilizing, making stratified equilibria more likely to result. Similarly, if the economy starts from a stratified equilibrium, the response to factors that generally tend to encourage human-capital investments may be highly asymmetric across local communities. In particular, we find that the negative global spillover effect that occurs since skilled workers from different communities meet in a common labour market can potentially overturn the positive incentive effects of e.g. a higher wage return to education and/or lower costs of education, potentially leading to increased differences in education and lower intergenerational social mobility. On the other hand, a policy that promotes equal opportunities in the labour market (strengthening "meritocracy") can enhance both efficiency and social mobility.

The paper is organized as follows. Section II describes the model. Section III characterizes steady-state equilibria. Section IV looks at the sustainability of non-stratified equilibrium. Section V takes a closer look at stratified equilibria in terms of their comparative statics properties, intergenerational social mobility, and efficiency. Finally, Section VI discusses the findings.

II The Model

Time is continuous and the horizon is infinite. An individual's lifetime has two phases: a working-life phase of length T and retirement phase (the length of which is irrelevant). Each individual belongs to a family characterized by overlapping generations: when an individual reaches retirement, her offspring enters working life. During working life the individual works and saves her earnings (consumption occurs only in retirement); her objective is to maximize total expected discounted earnings. There is a single numeraire consumption good in the economy.

A frequent succession of cohorts are born into the economy, each cohort being a continuum of unit size. A new cohort is born every $T/\mathcal{N}$ period, where $\mathcal{N}$ is a large natural number. Hence there is a total measure of $\mathcal{N}$ families, each with one member in the labour market. (Every family has a member only in each $\mathcal{N}$th cohort.) We assume that there are two exogenously given intrinsically identical locations, each containing $\mathcal{N}/2$ single-family houses; this will imply that each generation is split equally between the two locations. Workers may move at the point in time when they retire only, and this location choice determines the social environment of their offspring. Since this is the channel between parents and offspring that is the focus of the current paper we adopt the modeling strategy of making this is the only channel in that we assume that there is no bequests, no other parental investments, and no transmission of ability.

There are two types of jobs in the economy: "good" jobs and "bad" jobs. While good jobs can only be filled by skilled workers, bad jobs can be filled by any worker. Good jobs, are however, effectively rationed due to an incentive problem (described in detail below). A young individual, before entering the labour market, must decide whether or not to acquire skills. The cost of education has two components: a financial cost (e.g. a "tuition fee") plus an effort cost. There is a perfectly functioning credit market on which an agent can borrow to pay the tuition fee. The individuals, however, vary in "aptitude" or "talent", making the effort cost vary in the population. Due to the perfect credit market, we can combine the financial cost and effort cost into a single measure of an individual's idiosyncratic cost of education θ. We assume that θ is independent across identical individuals in the population, drawn from a distribution G with support $[\underline{\theta}, \overline{\theta}]$. In particular, talent is *not* transmitted from parent to offspring, and parents make their location choices unaware of their offspring's cost of education.

While education takes place locally within each community, workers meet in a common labour market; this creates scope for global (or "inter-community") spillovers. Global spillovers occur through the process by which information about job-openings is disseminated. A fraction of the attractive good jobs are allocated through connections. A skilled job-seeker may hear about a job-opening from another skilled "neighbor" (as described below), making the individual's expected returns to education depend in part on the composition of her local community. An individual cannot, however, choose her local community; this is determined by her parent's locational choice. Next we describe various components of the model in greater detail, starting with a description of the labour market.

Job Creation and Job-Finding Rates

A worker's productivity in a good job (for which skills are necessary) is constant equal to w_H while her productivity in an bad job is w_L (irrespectively of her skills); $w_H > w_L$. Bad jobs are always available to any worker who wants one; firms offering these jobs act competitively and pay the competitive wage w_L. Good jobs on the other hand may, due to an incentive problem, not be immediately available. There is on-the-job search in the sense that a skilled worker can work in a bad job while searching.[6] Once a worker finds a good job there are no natural separations so she can continue to work in that job until retirement. There is a separate job market for each cohort.

Following Saint-Paul (2001) we assume that a worker in a good job can try to access a "stealing technology". She might be caught trying, but if she is successful she can steal an amount h (at no risk) from the firm at every moment until retirement. A worker belonging to a given cohort is characterized by an age $t \in [0, T]$; we will consider a given cohort since good jobs will be age specific. We want to find the job-creation rate *at each age* that is compatible with a worker in a good job not misbehaving; under the assumption of free entry by firms, this will be the equilibrium job-creation rate, along with the competitive wage w_H.[7]

It is useful to derive value functions measuring expected discounted future earnings. Let $V^N(\tau)$ denote the value of being employed in a good job *not trying* to steal with $\tau \equiv T - t$ time left in the labour market. Trivially, for all $\tau \in [0, T]$,

$$V^N(\tau) = \frac{w_H}{r}\left(1 - e^{-r\tau}\right), \tag{1}$$

where r is the interest rate. Let $V^T(\tau)$ denote the value of *trying* to access the stealing technology with τ time left in the labour market. $V^T(\cdot)$ satisfies the asset equation

$$rV^T(\tau) = w_H + q\left[V^S(\tau) - V^T(\tau)\right] + p\left[U(\tau) - V^T(\tau)\right] - \dot{V}^T(\tau), \tag{2}$$

[6] The current model is in this sense similar to the "dual" labour market model in Bulow and Summers (1986).

[7] See e.g. Shapiro and Stiglitz (1984).

for all $\tau \in [0, T]$ where $V^S(\tau)$ is the value of having access to the *stealing* technology, $U(\tau)$ is the value of being fired, q is the probability of accessing the stealing technology, and p is the probability of being caught trying. $V^S(\tau)$ is similar to $V^N(\tau)$ (with $w_H + h$ in place of w_H) while $U(\tau)$ satisfies the asset equation

$$rU(\tau) = w_L + a(\tau)\left[V^N(\tau) - U(\tau)\right] - \dot{U}(\tau), \tag{3}$$

where $a(\tau)$ is the rate at which the worker finds another good job. The equilibrium job-creation adjusts so as to keep the "no-trying condition" $V^N(\tau) \geq V^T(\tau)$ satisfied with equality for all $\tau > 0$. In order for some job creation to take place we assume that $\Delta w \equiv w_H - w_L > qh/p$. Substituting for $V^S(\cdot)$ and $V^T(\cdot)$ in (2) gives

$$U(\tau) = \left[w_H - \frac{q}{p}h\right]\frac{1}{r}\left(1 - e^{-r\tau}\right). \tag{4}$$

Equation (3) can then be used to solve for the job-finding rate, $a(\tau) = rK/\left(1 - e^{-r\tau}\right)$, where

$$K \equiv \frac{p}{qh}\left(\Delta w - \frac{q}{p}h\right) = \frac{p}{q}\frac{\Delta w}{h} - 1 > 0. \tag{5}$$

The constant K measures inversely the labour-market distortion: the larger is K the more rapid is job-creation. Naturally, K is decreasing in h and q, and increasing in p and Δw.

Since each agent's lifetime is finite we can focus on the limiting case where there is no discounting. The job-finding rate at age t, denoted $a(t)$, then becomes[8]

$$a(t) = \frac{K}{T - t} \quad \text{for all } t \in [0, T). \tag{6}$$

Let $F(\cdot)$ denote the probability of finding a good job at age t or sooner associated with the hazard rate $a(\cdot)$; this is easily seen to be[9]

$$F(t) = 1 - \left(\frac{T - t}{T}\right)^K \quad \text{for all } t \in [0, T]. \tag{7}$$

The incentive problem determines the rehiring rate for a *job-loser* (i.e. a worker who has been fired after being caught misbehaving).[10] Thus we need to relate this rehiring rate to the hiring rates of "first-job-seekers". The latter rates may however be community-specific. Hence

[8]Note that we are abusing the notation slightly here by "reversing time"; $a(t)$ is obtained from $a(\tau)$ as $a(t) \equiv \lim_{r \to 0}\left[rK/\left(1 - e^{-r(T-t)}\right)\right]$.

[9]Note that $F(t)$ goes to unity as t goes to T, indicating that the worker will, with probability one, find a good job before reaching retirement.

[10]This is a common, but often implicit, feature of efficiency-wage model of unemployment.

let $N_j(t)$ denote the number of skilled job-seekers of age t in location j, and let $a_j(t)$ denote the rate at which these workers find good jobs. The labour market distortion creates effective "job rationing" when a job-loser cannot easily be distinguished from initial job-seekers. In order to capture the spirit of this we assume that a job-loser becomes indistinguishable from an "average" first-job seeker in his cohort in the sense that her rehiring rate equals the average hiring rate among the current first-job seekers.[11]

Assumption 1 *For all $t \in [0, T]$,*

$$a(t) = \frac{\sum_{j=1,2} N_j(t) a_j(t)}{\sum_{j=1,2} N_j(t)}.$$

Using this assumption we can show that the number of good jobs created, at any age t, depends only on the number of skilled workers in the cohort, not on their composition in terms of social background.

Proposition 1 *The number of good jobs created for workers of age t, denoted $M(t)$, is, for all $t \in [0, T)$, proportional to the number of skilled workers in the relevant cohort.*

The creation of good jobs for a given cohort is thus spread out over the entire period in which that cohort is in the labour market. In particular, the time of creation of a typical good job (as measured from the date at which the cohort enters the labour market) is a random variable drawn from the distribution $F(\cdot)$. Note that this implies an expected time of creation equal to $\bar{t}_0 \equiv T/(1+K)$.

Social Environment and Word-of-Mouth Communication

Each worker's local environment is determined by the residential choice of her parent. In particular, each community is characterized by a social environment, measured in terms of the skills of the workers living in that community. Let γ_j, $j = 1, 2$, denote the fraction of workers (of working age) in location j who are skilled. We will be focusing on steady-state equilibria and will hence treat γ_1 and γ_2 as constant over time.

One aspect of labour-market institutions is how information about job-openings is disseminated. To explore this we will assume that information about (a fraction of) the good jobs

[11] This implies that all job-losers become identical independently of their backgrounds. This simplifies the analysis in that the incentive condition will be the same for all workers in good jobs. Of course, in equilibrium, there will be no job-losers.

created is spread through word-of-mouth communication among skilled workers. We will show here that, under some simplifying assumptions, this leads to local job-finding hazards for good jobs that stand in constant proportion to each other.[12]

To do this we need to impose some spatial structure on the two locations. Recall that each location consists of $\mathcal{N}/2$ single-family houses. All houses are assumed to be symmetrically distributed in the sense that each family has n immediate "neighbors". Assuming that each family's neighbors are representative of the community, each skilled job-seeker in community j then has $n\gamma_j$ skilled neighbors. In order to derive the implications of the word-of-mouth communication process, we make the following assumptions:

Assumption 2 *(Word-of-mouth communication)*

1. *Information about good jobs can only be passed on between immediate neighbors.*
2. *There is no "relay" of information: information can only be passed on once.*
3. *Each skilled worker in the economy (whether currently in a good job or not) is equally likely to hear about any good job that is created.*
4. *Since vacancies are age-specific, the probability that a skilled worker who is first to hear about a specific vacancy will need it for herself, or will have more than one relevant neighbor to pass it on to, is negligible.*

The number of good jobs created for age-t skilled job seekers is $M(t)$. Assume for now that all good jobs are allocated through word-of-mouth communication. Consider then the probability that a specific job-seekers, of age t, in location j gets one of these $M(t)$ jobs. Note that $\sum_{i=1,2} N_i(t)\, n\gamma_i$ is the total number of skilled individuals who know *some* age-t job-seeking skilled worker; each of these individuals is equally likely to be the one who passes the information about any one of the $M(t)$ job-openings on to a worker who finally fills that specific vacancy. But the age-t job-seeker in location j knows $n\gamma_j$ of the potential information carriers; hence her job-finding hazard is simply

$$a_j(t) = \frac{n\gamma_j M(t)}{n\sum_{i=1,2} N_i(t)\,\gamma_i}. \tag{8}$$

[12] Our model of informal dissemination of information about job-openings can be viewed as a highly simplified version of the model in Calvo-Armengol and Zenou (2001).

The important property to note is that the local hazards stand in constant proportion to each other: for all $t \geq 0$,

$$\frac{a_1(t)}{a_2(t)} = \frac{\gamma_1}{\gamma_2}. \tag{9}$$

The relevant measure of the local social environment is thus the proportion of skilled workers γ_j in the location, and what matters for the allocation of good jobs is the relative proportion γ_1/γ_2. Intuitively, when there are more skilled workers in community 1 than in community 2, skilled worker from community 1 tend to find good jobs faster then their community-2 counterparts.

In order to be able to explore the role of labour market institutions we will however assume that some good jobs are formally advertised. While the chance of a worker getting an informally advertised job depends on her (relative) social environment as described above, all skilled job-seekers have an equal chance of getting a formally advertised job. To capture this we include a parameter $\upsilon \in [0, 1]$ to measure how many good jobs are allocated informally and generalize (9) to

$$\frac{a_1(t)}{a_2(t)} = (1 - \upsilon) + \upsilon\frac{\gamma_1}{\gamma_2} \equiv \varphi. \tag{10}$$

The special case $\upsilon = 1$ is the case where all good jobs are informally allocated, while $\upsilon = 0$ is the case where all good jobs are formally allocated.

Local Complementarities in Human Capital Investments and Educational Choices

A social environment may impact both on the cost of and the return to education. The mechanism described above, where an agent with more skilled neighbors is more likely to hear about job-openings, works through the expected return to education. Note that this is a *global* (or "inter-community") spillover effect in the sense that a worker in one community is affected by the social environment in the adjacent community. We will also allow the social environment to affect the agent's cost of education through a *local* human-capital spillover effect: in each community, the more people acquire skills, the easier/less costly it is to invest in education. There are several reasons why there might be local complementarities in human capital investments: fiscal externalities, scale economies, peer group effects etc. Though this local (or "intra-community") spillover effect is not essential to the main results regarding efficiency properties of symmetric communities and the importance of formal versus word-of-mouth communication, we include it since it interacts with the global spillover effect in an interesting way.

We want, for example, to consider whether local spillovers in human capital investments tend to stabilize or destabilize a non-stratified outcome in the presence of global spillovers.

Hence we assume that an agent's effective cost of education is lower when a larger fraction of the individuals in his community also invest in education. Thus let the effective cost of education to an individual from community j, with idiosyncratic cost θ, be $\theta\pi(\gamma_j)$, where γ_j is the fraction of workers from the same community (and cohort) who also invest in education; the function $\pi(\cdot)$ is assumed to be continuous and decreasing with $\pi(0) = 1$ and $\pi(1) > 0$.

Let us now consider the incentives for investing in education. Each young worker takes his location, the social environments, and the behaviour of everyone else as given. First we need to verify that the social environments, γ_1 and γ_2, generate well-defined community-specific job-finding prospects. To that end, let $F_j(t)$, $t \in [0, T]$ denote the probability that a skilled worker in location j finds a good job prior to age t.

Lemma 2 *Given that, in every cohort, a fraction $\gamma_j > 0$ of the workers in community j, $j = 1, 2$, invest in education, $F_1(\cdot)$ and $F_2(\cdot)$, are uniquely identified, and satisfy $[1 - F_1(t)] = [1 - F_2(t)]^{\varphi}$ for all $t \in [0, T]$. If $\gamma_j = 0$, then $F_k(\cdot) = F(\cdot)$.*

This makes clear the role of the social environments for job-finding prospects: if there are more skilled workers in community j than in community k, that is, if $\gamma_j > \gamma_k$ (and some good jobs are informally allocated), then $F_k(\cdot) < F(\cdot) < F_j(\cdot)$, implying that $F_k(\cdot)$ stochastically dominates $F_j(\cdot)$.[13] Thus a skilled worker from community k is disadvantaged relative to a worker from community j; since she has relatively few skilled workers in her local environment she can expect to wait longer to find a good job. Let $\bar{t}_j = \bar{t}_j(\gamma_j, \gamma_k)$ denote the expected time a skilled worker in community j must wait before she finds a good job; it then follows that $\gamma_j \geq \gamma_k$ implies $\bar{t}_j \leq \bar{t}_k$.

Consider then the educational choice of a worker in community j with idiosyncratic cost θ (and hence effective cost $\pi(\gamma_j)\theta$). If she remains unskilled her lifetime earnings will be Tw_L; if, on the other hand, she invests in education, her expected lifetime earnings will be $Tw_H - (w_H - w_L)\bar{t}_j$. Hence, she will make the investment if and only if:

$$\theta < \theta_j \equiv \frac{\Delta w}{\pi(\gamma_j)} \left(T - \bar{t}_j(\gamma_j, \gamma_k)\right). \tag{11}$$

If more workers invest in education in community j than in community k, then an individual young worker in community j has a stronger incentive to invest in education than an

[13] The fact that $F(\cdot)$ is "in between" follows from (A4) in the Appendix.

identical young worker in the other community. Note that this is true even if there are no local complementarities in education. While any local complementarity implies that a young worker in community j has a lower *cost* of education, the social network effect implies that he also has a higher expected *return* than his community-k counterpart.

Prior to realization of a worker's idiosyncratic education cost θ, her expected lifetime income (net of any idiosyncratic educational cost incurred), denoted z, is community-specific:

$$z_j \equiv \int_{\underline{\theta}}^{\overline{\theta}} \max\left\{Tw_L, Tw_H - \Delta w \overline{t}_j\left(\gamma_j, \gamma_k\right) - \theta\pi\left(\gamma_j\right)\right\} g\left(\theta\right) d\theta, \qquad j = 1, 2. \tag{12}$$

From the above discussion it follows immediately that $\gamma_j > \gamma_k$ implies that $z_j > z_k$.

Locational Choices

Whenever the proportions of educated workers differ across communities, the community with the higher proportion will constitute a more favorable social environment. This creates the potential for a self-reinforcing feedback where the superiority of one community persists due to the education incentives coming with a better environment. If such differences persist, they will have implications for intergenerational social mobility and the exact link will depend on how individuals sort themselves into communities.

Given that the communities may differ in social environment, and given that social environment is a key determinant of a child's prospects, one natural way in which sorting will occur is if parents care about the future prospects of their offspring, as measured e.g. by the net expected lifetime earnings, z. A more favorable location will then naturally command a higher rental price. Moreover, any mild complementarity between own consumption and concern for the offspring will induce stratification by income: parents with realized incomes in the upper half of the aggregate income distribution will outbid the parents with incomes in the lower half for the houses in the attractive location. For simplicity we assume that parents are only mildly altruistic. The rental price difference will then be small (implying that we are justified in focusing on maximization of lifetime income as an individual's objective) and parents will not want to leave any bequests.[14]

[14] Assume e.g. that a parent's preferences over own consumption and the child's prospects are cz^{α}, where $\alpha > 0$ but small. Let ρ_j be the rental price in area j, and suppose e.g. that $\gamma_1 > \gamma_2$. The locational choice of an individual with realized lifetime income y can be described as $j^*\left(y\right) = \arg\max_j \left\{\left(y - \rho_j\right) z_j^{\alpha}\right\}$ which is monotonic in y for any $\alpha > 0$. If α is small the equilibrium rental prices must be close, and moreover, the indirect

Local and Aggregate Income Distributions

Before turning to steady-state equilibria, we show that a time-invariant skill distribution maps into two local (lifetime) income distributions, Φ_1 and Φ_2, and a global income distribution Φ. Thus suppose that a fraction γ_j of the workers in location j invest in education (in every cohort). The support of Φ_j is $[Tw_L, Tw_H]$. For every possible realization of lifetime earnings y in this support, $\Phi_j(y)$ has two parts: the probability that a randomly chosen agent in community j remains unskilled (which has probability $1-\gamma_j$), and the probability that the agent invests in education but earns less than y. Calculating the conditional probability of the latter event using the local distribution of job-finding dates $F_j(\cdot)$, we obtain that the distribution of lifetime earnings among the workers in community j is

$$\Phi_j(y) = 1 - \gamma_j F_j\left(\frac{Tw_H - y}{\Delta w}\right). \tag{13}$$

Since half of each cohort is born into each community, the economy-wide income distribution, denoted Φ, is simply $\Phi = (\Phi_1 + \Phi_2)/2$. Moreover, since F_1 and F_2 are uniquely determined by γ_1 and γ_2 (Lemma 2) so are the income distributions, Φ_1, Φ_2, and Φ.

III Steady-State Equilibrium

So far we have taken the fractions of skilled workers in the two communities as given and shown how γ_1 and γ_2 determine (i) job creation, (ii) local job-finding rates, and (iii) local and aggregate income distributions. However, we have not required that γ_1 and γ_2 be consistent with rational education choices. A worker in community j will only invest in education if he is talented enough so that $\theta < \theta_j$, defined in (11). The fraction of agents in community j acquiring skills will then be $\gamma_j = G(\theta_j)$. Rationality and consistency can therefore succinctly be summarized in the following two equilibrium conditions:

$$\gamma_j = G\left(\frac{\Delta w}{\pi(\gamma_j)}\left(T - \bar{t}_j(\gamma_j, \gamma_k)\right)\right), \qquad j,k = 1,2,\ j \neq k. \tag{14}$$

In words, (14) says that the educational choices of the agents in each community should be consistent with the incentives created by the social environments – in both communities – that

utility of lifetime income, $\left(y - \rho_{j^*(y)}\right) z^{\alpha}_{j^*(y)}$ becomes effectively linear which is consistent with assumption that the agents maximize lifetime earnings during working life. This story, however, depends on parents not knowing the education cost of their offspring at the point in time when they choose location.

those educational choices generate. From now on we will refer to a pair (γ_1, γ_2) that solves (14) for $j = 1$ and $j = 2$ simultaneously as a *steady-state equilibrium*.

It is useful to define a *conditional community equilibrium* as a γ_j that solves (14) for a given $\gamma_k \in [0, 1]$. Uniqueness is implied by the following "within-group stability condition" which we take to hold throughout: for all $\gamma_j \in [0, 1]$,

$$\pi'(\gamma_j)\theta(\gamma_j) + \frac{\pi(\gamma_j)}{g(\theta(\gamma_j))} + \Delta w \frac{\partial \bar{t}_j}{\partial \gamma_j} > 0, \tag{15}$$

where $\theta(\gamma_j) = \theta_j \equiv G^{-1}(\gamma_j)$ is the cut-off education cost of the last individual to invest in education.[15] We can then write the conditional community equilibrium as a continuous function $\gamma_j(\gamma_k)$ and characterize a steady-state equilibrium as a pair (γ_1, γ_2) for which $\gamma_j = \gamma_j(\gamma_k)$ for $j, k = 1, 2$ simultaneously.

There may exist two types of steady-state equilibria: symmetric (or "non-stratified") and asymmetric (or "stratified"). While asymmetric equilibria may or may not exist, there always exists exactly one symmetric equilibrium, henceforth denoted by $*$, the proportion of workers with an education in each community being γ^*. Moreover, the symmetric equilibrium is independent of the way information about good jobs is disseminated.

PROPOSITION 3 *There exists exactly one symmetric steady state equilibrium γ^*, which is independent of v, and which has $\bar{t}_j = \bar{t}_0 \equiv T/(1 + K)$ for $j = 1, 2$.*

In the symmetric equilibrium the two communities offer exactly the same social environment; hence, individuals in both locations behave in the same way, and no parent is willing to pay more to live in one community than in the other. Another feature of the symmetric equilibrium is that there is perfect intergenerational mobility: there is zero correlation between the lifetime income of a child and that of her parent.[16]

However, non-stratified communities is not the only possibility. Intuitively, there can be stable outcomes where one community offers a more favorable social environment and, as a consequence, young people are more prone to invest in education there, thus perpetuating the

[15] The condition guarantees that the derivative of the right-hand side of (14) not exceeds one, which corresponds to a local expansion of education not being self-reinforcing absent adjustment in the other community. The condition involves local and the global spillover effects, and rules out the local externality being too strong.

[16] To see this note that since the communities are identical the local income distributions are the same $\Phi_j = \Phi$, $j = 1, 2$; moreover, the child's and the parent's incomes are independent draws from Φ. Since independence implies zero covariance, the incomes are uncorrelated.

difference in the social environments. The scope for stratification stems from the fact that the workers from the two communities meet in a common labour market; global spillover effects arise since the local social environments in *both* communities affect the prospects of young workers in each of the two communities. To gain further insight into the global spillover, consider how the local "waiting times" $\bar{t}_1$ and $\bar{t}_2$ are affected by marginal increases in γ_1 and γ_2. An increase in γ_j has two effects. First, it improves the set of connections that a young worker in community j can use to find a good job; this effect comes at the expense of the other community and hence tends to reduce $\bar{t}_j$ and increase $\bar{t}_k$. Second, when γ_j increases, the total number of skilled workers in the economy increases; the firms respond to this by creating equally many new jobs. However, the new jobs are created with an average waiting time of $\bar{t}_0$; if human capital investments are increasing in the relatively disadvantaged community – i.e. if initially $\gamma_j < \gamma_k$ – the expected waiting time for the newly created jobs is shorter than $\bar{t}_j$, allowing *all* skilled workers to find good jobs, on average, faster, thus reducing both $\bar{t}_j$ and $\bar{t}_k$. If, on the other hand, human-capital investments are increasing in the already advantaged community – i.e. if initially $\gamma_j > \gamma_k$ – this effect goes in the opposite direction, increasing both $\bar{t}_j$ and $\bar{t}_k$. The thing to note is that the effect of expanding human-capital investments in either community on the disadvantaged community is always unambiguous.

LEMMA 4 *If* $\gamma_j \leq \gamma_k$, *then* $\partial \bar{t}_j / \partial \gamma_j < 0$ *and* $\partial \bar{t}_j / \partial \gamma_k > 0$, *while if* $\gamma_j > \gamma_k$ *both effects are ambiguous.*

The effect of the global spillovers is also easy to see from the conditional community equilibria, $\gamma_j(\cdot)$. Suppose e.g. that fewer (more) than γ^* workers invest in education in community k; the rational response is then for more (fewer) than γ^* workers to invest in education in community j.[17] See Figure 1. Note that $\gamma_j(0) = \gamma^*$ since, when no one invests in human capital in community k, the expected waiting time for skilled workers in community j is necessarily $\bar{t}_0$.

Figure 1 here

As the figure makes clear, the number of steady-state equilibria will always be odd: if (γ_1, γ_2) is an equilibrium, its mirror image is also an equilibrium. Hence we can focus on equilibria where $\gamma_1 \geq \gamma_2$. Moreover, even though there can be more than three equilibria, these cases are both implausible and less interesting; hence from now on we restrict attention to cases where there are at most three equilibria.

[17] This follows from Lemma 5, and from noting that $\gamma_k \neq \gamma^*$, $\gamma_k \neq 0$ implies $\gamma_j(\gamma_k) \neq \gamma^*$.

Stability

Our notion of stability is the standard one based on the reaction functions, *viz.* the one requiring that a myopic adjustment process with the γ's being adjusted alternately converge to the equilibrium.[18] The following Lemma is straightforward.

LEMMA 5 *A symmetric equilibrium is locally stable if and only if*

$$\frac{\partial \gamma_j}{\partial \gamma_k} = \frac{-\Delta w \partial \bar{t}_j / \partial \gamma_k}{\pi'(\gamma_j)\,\theta_j + \pi(\gamma_j)\,/g(\theta_j) + \Delta w \partial \bar{t}_j / \partial \gamma_j} > -1 \tag{16}$$

at $\gamma_j = \gamma_k = \gamma^*$.

It is worth noting that $\gamma_j(\cdot)$ slopes downward at the symmetric equilibrium since the numerator is positive by Lemma 4, and the denominator is positive by the within-group stability condition. Note that, by the expression given in (16) and using Lemma 4, $\gamma_j(\cdot)$ slopes downward when $\gamma_j < \gamma_k$; in particular, $\gamma_j'(\gamma_k) < 0$ when $\gamma_k > \gamma^*$. This, combined with the fact that $\gamma_j(0) = \gamma_j(\gamma^*) = \gamma^*$ and the previously noted fact that the number of equilibria is odd, makes clear that there are precisely two cases when the number of equilibria is a most three:

- The symmetric equilibrium is stable, and it is the unique equilibrium.
- There are two asymmetric equilibria in addition to the symmetric one; the asymmetric equilibria are stable while the symmetric equilibrium is unstable.

Efficiency

Are any of the steady-state equilibria ever efficient? And if not, what are the sources of inefficiency? Since the individuals' preferences are linear in consumption, it is natural to use

[18] The out-of-steady-state dynamics of the current model are clearly very complicated. Explicit dynamics can be studied in simplified versions; suppose e.g. that only two generations of workers – young and old – are present in the labour market at the same time and that old workers can provide young workers with information about good jobs (while young workers hear about formally advertised jobs directly). When studying out-of-steady-state dynamics one needs to consider the (unique) equilibrium behaviour of one cohort conditional on the behaviour of the previous cohort; this determines the evolution of the economy (which is deterministic). The evolution then determines steady state equilibria. The simplified model, with explicit dynamics, shares the key properties of the current model: there is exactly one symmetric steady state equilibrium, and there may be other asymmetric equilibria. If the symmetric equilibrium is the unique steady state equilibrium, it is stable; if there are three steady state equilibria, the asymmetric ones are stable, while the symmetric one is unstable.

total surplus, defined as total output minus aggregate education cost, as the efficiency criterion. What then does an efficient allocation look like? And in particular, will having non-identical communities ever be superior to a symmetric outcome?

The answer to the last question is no. The reason for this is simple. Since the creation of good jobs is proportional to the total number of skilled workers in each cohort, aggregate production is linear in the aggregate number of skilled workers: the aggregate output of a cohort is equal to $\overline{\gamma}\left[Tw_H - \Delta w\overline{t}_0\right] + (1-\overline{\gamma})\,Tw_L$, where $\overline{\gamma} = (\gamma_1 + \gamma_2)/2$ is the total number of skilled workers in that cohort. This implies that there are no gains in terms of output to having asymmetric communities.

Moreover, since any potential spillovers in education are only local, the cost of educating a given number of workers in community j is independent of the number of workers acquiring skills in community k.

Combining these two observations immediately leads us to conclude that there can be no efficiency gains from having asymmetric communities.

PROPOSITION 6 *There is always an efficient allocation in which the communities are symmetric.*

This result does not require any convexity assumption. Convexity of the cost of educating γ workers within a community would ensure the existence of a unique efficient allocation, which then by Proposition 6 is symmetric: denote the cost of educating a fraction γ of a community's cohort (measured per community member in the relevant cohort) as

$$C(\gamma) \equiv \pi(\gamma)\int_{\underline{\theta}}^{\theta(\gamma)} \theta g(\theta)\, d\theta, \tag{17}$$

where as before $\theta(\gamma) = G^{-1}(\gamma)$. This cost is convex if

$$C''(\gamma) = \pi''(\gamma)\int_{\underline{\theta}}^{\theta(\gamma)} \theta g(\theta)\, d\theta + 2\pi'(\theta)\,\theta(\gamma) + \frac{\pi(\gamma)}{g(\theta(\gamma))} > 0. \tag{18}$$

Here we see that two factors in particular tend to make $C(\cdot)$ convex. Decreasing returns to the local spillover in education – $\pi'' > 0$ being large relative to absolute value of $\pi' < 0$ – is one such factor. The other is when there is large variability in talent, implying that the density $g(\cdot)$ is small; generally, the larger is the dispersion in talent, the more likely is it that the marginal cost of educating additional workers will be increasing. Note that if there are no local complementarities in education (i.e. $\pi(\gamma) = 1$ for all γ) then $C(\cdot)$ is always strictly convex due to the variation is talent.

If there are local complementarities in education, then not even the symmetric equilibrium γ^* will be efficient. This is the standard externality that a worker does not take into account the effect of his education decision on the education costs for the other cohort members in his community. This is, however, the only inefficiency of the symmetric equilibrium.

A stratified equilibrium will generally involve a second inefficiency: given that $C(\cdot)$ is convex the aggregate cost of educating a total fraction $\overline{\gamma}$ of the workers in a cohort is minimized by educating the same fraction in both communities, $\gamma_1 = \gamma_2 = \overline{\gamma}$. Hence, on top of the fact that the number of skilled workers is generally wrong, in an asymmetric equilibrium the aggregate cost of educating fails to be minimized: the equilibrium number of skilled workers could be educated more cheaply by ensuring that the marginal students in the two communities are equally talented.[19]

IV Sustainability of a Non-Stratified Equilibrium

What factors contribute to making stratified equilibria less likely to be sustainable? Can we expect factors that promote human-capital acquisition also to promote social integration? A partial answer to this question can be obtained by considering in detail the stability of the symmetric equilibrium. We will show that, perhaps somewhat surprisingly, factors that generally encourage investments in human capital can be inherently destabilizing, thus contributing to making social stratification more likely.

In order to sort out the forces at work in determining the stability of the symmetric equilibrium – and thus the number of equilibria – it is useful to make the following simplifying assumptions.

ASSUMPTION 3 *The distribution of θ is uniform on $[0, A]$, for some $A > 0$, and the local externality of the number of educated workers in a location implies exponentially decaying education costs, i.e., $\pi'(\gamma)/\pi(\gamma) = -a$, for some constant $a \geq 0$.*

With the uniform distribution for θ the (constant) density is $g = 1/A$. The substance of assuming that θ is uniformly distributed is that we rule out effects stemming from the density function being increasing or decreasing locally.[20]

[19] When $C(\cdot)$ is not convex, an asymmetric allocation *may* be optimal – although not uniquely optimal – and an efficient allocation is not in general characterized by the first-order condition.

[20] As we have seen – for example, in the within-group stability condition (15) – a small value of the density at

Recall that the symmetric equilibrium is stable if and only if $\left|\partial\gamma_j/\partial\gamma_k\right| < 1$ (at γ^*). We will therefore say that a factor *destabilizes* (*stabilizes*) the symmetric equilibrium if it, ceteris paribus, increases (decreases) $\left|\partial\gamma_j/\partial\gamma_k\right|$.

With a uniform distribution and no local spillovers (i.e. $a = 0$), the stability of the symmetric equilibrium is independent of γ^*, the number of workers acquiring education. However, when there are local complementarities in education, then the scale of the symmetric equilibrium matters; in particular, the larger is γ^* the less likely it is that the symmetric equilibrium is stable – an increase in γ^* directly increases $\left|\partial\gamma_j/\partial\gamma_k\right|$. This raises the possibility that some factors may affect the stability of the symmetric equilibrium *only* through its effect on its scale; as we will see shortly, this is true for the wage return, Δw, which increases γ^*. We can thus conclude, for example, that an increase in the wage return will affect stability only if there are local spillovers, and will then affect it negatively (see Proposition 7 below).

We will label effects that occur via γ^* as "scale effects", and refer to the direct effect on $\left|\partial\gamma_j/\partial\gamma_k\right|$ as the "sensitivity effect" (reflecting that the impact is via the responsiveness of γ_j to γ_k and vice versa). The factors that we will consider are:

1. A reduction in education cost, represented by an increase of the (constant) density g.[21]

2. An increase in the wage return, represented by an increase in Δw (holding the speed of creation of good jobs fixed[22]); note that since the wages reflect underlying technologies, an increase in Δw may be interpreted as a "skill-biased" technological change.

3. A decrease in the labour market distortion, represented by an increase in K, which leads to faster creation of good jobs.

4. A strengthening of the local education spillovers, represented by an increase in a.

5. An increase in the fraction of jobs that are allocated through social networks, represented by an increase in v.

---

the equilibrium serves to stabilize the equilibrium; similarly, it turns out that a *locally increasing* density leads to any increase in the number of workers acquiring education having a destabilizing effect.

[21] I.e., we simply scale down the interval $[0, A]$ over which θ is uniformly distributed; this corresponds to each worker's education cost being reduced proportionally by a factor common to all workers.

[22] Note that Δw enters as a determinant of K, which determines the speed of creation of good jobs. Thus we ignore this channel and treat an increase in K separately.

We can summarize the unambiguous effects as follows.

PROPOSITION 7 *Stronger local spillovers (a) and larger reliance on word-of-mouth communication (v) are always destabilizing; when $a > 0$ reduced education costs (an increase in g) and a higher wage return (Δw for fixed K) are destabilizing too.*

That the importance of informal job allocation, v, is destabilizing is due solely to the sensitivity effect (since v does not affect the scale γ^*).[23] The effect of the wage return Δw is due solely to the scale effect (which is why it only occurs when there are local spillovers in education). The effect of reduced education costs emanates from both channels; both effects vanish, however, when $a = 0$. The only ambiguous effect is a decrease in the labour market distortions; an increase in K leads to faster job creation and an increase in the scale γ^*. But on the other hand, it also implies shorter waiting times on average, thus indirectly reducing the importance of social networks and, hence, sensitivity. Hence, K the effect of generally ambiguous, but is stabilizing when a is zero (or sufficiently small).

The properties demonstrated underline a recurrent observation, *viz.* the potential for policies to have ambiguous effects. In particular, we have seen that education is generally underprovided in equilibrium, which calls for policies that promote education; such policies may, however, upset a symmetric equilibrium leading to inefficient – and otherwise undesirable – segregation.

V Effects on a Stratified Equilibrium

The analysis of the stability of the symmetric equilibrium provided us with some insight into what factors avert social stratification. We now proceed to consider whether the same factors that permit stratified equilibria also make such equilibria in some sense more polarized. We maintain Assumption 3 since it offers us a simple parameterization of education costs and of local education spillovers. Our starting point is thus the existence of a stratified equilibrium with $\gamma_1 > \gamma_2$, and our first concern is how various factors affect investments in human capital as measured by γ_1 and γ_2.

As in the case of stability, we will be interested in the effect of a number of factors: a reduction in education costs (represented by an increase in g), an increase in the wage return to skill Δw (ignoring the effect via job creation), a decrease in labour market distortions (represented

[23] Moreover, the conclusion concerning v does not require Assumption 3.

by an increase in K) leading to faster creation of good jobs, a strengthening of local education spillovers (represented by an increase in a), and an increase in the fraction of good jobs that are allocated through word-of-mouth communication (represented by an increase in υ). It is useful to derive a general expression for the equilibrium effect on γ_j of a generic parameter α, $\alpha = g, \Delta w, K, a, \upsilon$. Recall that (14) defines the locus of conditional community equilibria $\gamma_j(\cdot)$. Moreover, we know that, $\gamma_1 > \gamma_2$ implies $\gamma_2'(\gamma_1) < 0$ while $\gamma_1'(\gamma_2)$ has an ambiguous sign (see Figure 1). The equilibrium effects on participation in education in the two communities depend partly on the direct effects on education incentives and partly on global spillovers. Define the partial or "direct" effect of α on γ_j, denoted $\partial\gamma_j/\partial\alpha$, as the effect on γ_j that would have obtained had γ_k remained unchanged (obtained by totally differentiating (14) while holding γ_k fixed). In terms of Figure 1, a positive direct effect, $\partial\gamma_j/\partial\alpha > 0$, simply means that $\gamma_j(\cdot)$ shifts upwards locally. Then denote the total effect on the steady-state equilibrium value of γ_j – which includes global spillover effects – by $d\gamma_j/d\alpha$. It is then straightforward to show that

$$\frac{d\gamma_j}{d\alpha} = \frac{1}{D}\left(\frac{\partial\gamma_j}{\partial\alpha} + \gamma_j'(\gamma_k)\frac{\partial\gamma_k}{\partial\alpha}\right), \quad j,k = 1,2,\ j \neq k, \tag{19}$$

where $D > 0$ (given that the equilibrium is locally stable).

We can partition the parameters into two distinct groups according to their direct effects.

The factors that have strictly positive direct effects include (i) a reduction in education costs, (ii) an increase in the wage return Δw, (iii) a reduction in labour market distortions leading to faster creation of good jobs, or (iv) stronger local spillover effects in education. These factors thus directly encourage education: $\partial\gamma_j/\partial\alpha > 0$ for these factors. However, this does not rule out asymmetric responses across the two communities. Consider first community 2. For this community the positive direct effect is counteracted by a strictly negative global spillover effect from community 1 – the second term in the parenthesis in (19) is strictly negative.

This is due to the fact that an increase in γ_1 makes it more difficult for workers in community 2 to find good jobs quickly.

If the direct effect is relatively strong in the advantaged group, and the responsiveness in community 2 to an increase in γ_1 is fairly strong, then the direct effect may by substantially or completely outweighed by the negative spillover effect. It is interesting to note that the responsiveness, measured by the slope $\gamma_2'(\gamma_1)$, *arises fundamentally* from the global spillover effect due to the workers meeting in a common labour market, but is *amplified* by any local complementarities in education.

For the advantaged community 1, the direct positive effect is less likely to be counteracted by global spillover effects. Indeed, the global spillovers may even reinforce the direct effect for this group – the second term in (19) may be positive (see Lemma 4).

The reason is that increased education in community 2 leads to equally many good jobs being created, but at a faster rate than the workers in community 2 find jobs; this effect tends to allow the skilled workers from community 1 to find good jobs more quickly. Figure 2 illustrates a case where equilibrium education increases only in the advantaged community.

Figure 2 here

A change in the importance of social contacts in the job-finding process on the other hand has a markedly different effect. Suppose e.g. that the fraction of good jobs that are allocated through social networks, υ, is increased. The direct effect of this is to make it relatively more difficult for workers in the community 2 to find good jobs quickly; in particular, the increase in υ causes $\bar{t}_1$ to decrease and $\bar{t}_2$ to increase. This implies that direct effect $\partial\gamma_j/\partial\upsilon$ is negative for community 2 and positive for community 1. Since the global spillover effect affecting community 2 is then also strictly negative, the equilibrium effect is to reduce γ_2.

Hence we find that factors that generally encourage education have ambiguous effects on participation in education, with a marked possibility that the educational responses can be highly asymmetric across stratified communities. The only factor for which the comparative-statics effect is unambiguous is an increases in υ, the fraction of jobs that are allocated through word-of-mouth communication, on participation in education in the disadvantaged community. (The proof of the following Proposition gives the details of all the above-mentioned results.)

Proposition 8 *An increase in the fraction of good jobs that are allocated through social networks, υ, unambiguously decreases education in the disadvantaged community.*

The analysis highlights how social interaction can potentially account for aggregate educational responses that may be difficult to explain with an individualistic framework with atomistically optimizing agents, especially if credit constraints are unlikely to be binding heavily. For example, in the current framework factors that generally encourage human-capital accumulation may have the equilibrium effect of making participation in education more polarized across socio-economic groups.

Effects on Mobility and Efficiency

The effects of stratification are also reflected in the degree of intergenerational social mobility. A simple mobility measure in the current model is the fraction of families that switch from having above median income in one generation to having below median income in the next (or vice versa). This is of course also the fraction of families that switch community 1 to community 2. Denote this fraction by μ and note that it satisfies $\mu = \Phi_1(\widehat{y})$ where $\widehat{y}$ is the aggregate median income.

Since half of each cohort is born in to each community, $\widehat{y}$ must identically satisfy $\Phi_1(\widehat{y}) + \Phi_2(\widehat{y}) = 1$. Consider then the effect of a generic parameter $\alpha = g, \Delta w, K, a, \upsilon$ on mobility μ in a stratified equilibrium. Taking into account the impact of α on both the local income distributions and on the median income we obtain that

$$\frac{\partial \mu}{\partial \alpha} = \frac{\phi_2}{\phi_1 + \phi_2} \frac{\partial \Phi_1}{\partial \alpha} - \frac{\phi_1}{\phi_1 + \phi_2} \frac{\partial \Phi_2}{\partial \alpha}, \tag{20}$$

where ϕ_j is the density function of the local income distribution evaluated at $\widehat{y}$.

This expression shows that an unambiguous effect on mobility would obtain if the local income distributions Φ_1 and Φ_2 were to move in "opposite directions" (in the first-order stochastic-dominance sense). The above analysis then suggests that it is quite conceivable that factors that generally encourage education (e.g. a reduction in education costs, an increase in the wage return etc.) may have a negligible or even negative impact on mobility. As noted above, asymmetric responses with a stronger positive response in the already advantaged group is a marked possibility, which would then tend to reduce mobility and a stronger relation between parental income and participation in education.

One factor can, however, be expected to almost certainly increase measured mobility, *viz.* a reduction in the role of word-of-mouth communication. If a reduction in υ not only increases participation in education in the community 2 (as shown in Proposition 8) but also (weakly) decreases it in the community 1, the effect on measured intergenerational mobility will be strictly positive.[24] This illustrates that an understanding of cross-country differences in intergenerational mobility may require going beyond a comparison of educational systems to consider also the role of social networks and labour market institutions.

[24] To see this, recall that $\Phi_j(y)$ is given by (13). By applying the technique used in the proof of Lemma 4 one can show that $\partial F_1/\partial \upsilon > 0$ and $\partial F_2/\partial \upsilon < 0$. Then, using that $d\gamma_2/d\upsilon < 0$ and $d\gamma_1/d\upsilon > 0$, it follows that $\partial \Phi_1/\partial \upsilon < 0$ and $\partial \Phi_2/\partial \upsilon > 0$, whereby from (20), $\partial \mu/\partial \upsilon < 0$.

Finally a word on efficiency. Local spillovers imply that there are generally too few skilled workers; this suggests that education should be encouraged e.g. through a general policy with subsidies. However, the above analysis indicates that a general policy need not be unambiguously efficiency enhancing. To see this, suppose that it leads to an asymmetric response with a substantial increase in γ_1 but a modest response in γ_2. Due to the global spillover effect, skilled workers in community 1 find good jobs at rate that exceeds the rate at which such jobs are created. Thus, for this reason, their private benefit to education tends to exceed the social benefit. This of course simply reflects the global externality.[25]

VI Discussion

Empirical evidence suggests that intergenerational social mobility varies greatly across countries. Low social mobility reflects, among other things, the positive relationship between parental income and children's education. The most common explanation among economists is that family income matters for education because of credit constraints. A growing empirical literature suggests, however, that the link between parental income and children's education decisions may emerge due to more long-run factors affecting children's development and opportunities. It is for example frequently argued that rich and poor parents differ in the attitudes and norms they instill in their children; similarly, parental income may be correlated with the type of social networks that the children inherit.

Determining the mechanisms underlying the positive education-parental income relationship is of great importance in order to identify appropriate policies to promote human-capital investments and intergenerational social mobility. For example, policies designed to alleviate financial constraints – such as education subsidies and income transfer policies – may have unintended consequences if the link between parental income and education decision is, in fact, not due to credit constraints. This point has recently been stressed by several authors; Cameron and Heckman (1998), after finding empirically that long-run factors appear to be of primary importance, simulate the effects of a general increase in family income on college attendance. Their conclusion is that, while increasing resources available to parents is likely to increase enrollment,

[25] The effect of a generic parameter α on total surplus in a steady state equilibrium can be shown to be

$$\frac{1}{2}\sum_{j=1,2}\left[\left(\bar{t}_j-\bar{t}_0\right)\Delta w-\pi'\left(\gamma_j\right)\int_{\underline{\theta}}^{\theta(\gamma_j)}\theta g\left(\theta\right)d\theta\right]\frac{\partial\gamma_j}{\partial\alpha}$$

Note that $\bar{t}_1<\bar{t}_0$ and $\bar{t}_2>\bar{t}_0$ in a stratified equilibrium with $\gamma_1>\gamma_2$.

the students that are attracted into college have considerably lower ability than those already opting to attend. Similarly, Mulligan (1997), looking at intergenerational social mobility, finds no support for the existence of significant credit constraints and concludes that government policies designed to remove financial barriers to participation in education need not necessarily lead to higher mobility.

A different view on policy, based on the notion that more long-run factors are responsible for the education-parental income relationship, is that governments should work to change the structure of the economic environment in order to “level the playing field”, minimizing the effects of socioeconomic stratification and labour market segmentation.

The current paper has picked up on these ideas by formalizing one particular role played by social environments, namely the role played by social networks in job-finding processes. While the model is still a version of a parental investment model in the sense that rich parents can “invest in good neighbors” in order to improve the lifetime opportunities of their children, credit constraints do not play a role. The relation between parental income and education choices is thus indirect, emerging through the link between parental income and the social environment in which the children grow up.

The model provides one example of how, with this type of causal link, intuitions about the effects of policy or other shifts in the environment may be very different from those obtained in models based on credit constraints. We found, for example, that a general reduction in the cost of education or an increase in the returns to skill, could lead to highly asymmetric educational responses across socioeconomic groups. The model also contains one parameter which captures the notion of “leveling the playing field”; reducing the number of jobs that are allocated through word-of-mouth communication promotes equal opportunities in labour markets, and was found to generally promote both mobility and the efficient use of talent.

Appendix

Proof of Proposition 1. Note first that $a_i(t) = -\dot{N}_i(t)/N_i(t)$. Hence from (6) and Assumption 1 we then obtain $-\dot{N}(t)/N(t) = K/(T-t)$ for all t, where $N(t) = \sum_{i=1,2} N_i(t)$ is the aggregate number of skilled workers, of age t, looking for good jobs. Solving the differential equation yields

$$N(t) = \left(\frac{T-t}{T}\right)^{K} N(0) \tag{A1}$$

where $N(0)$ is the total number of workers in the cohort acquiring skills. Since all vacancies are filled immediately, $M(t) = -\dot{N}(t)$, whereby the result follows immediately. ■

Proof of Lemma 2. Let $\gamma_1, \gamma_2 > 0$ and recall that half of each cohort are born into each community. The number of agents in location j investing in education is then $N_j(0) = \gamma_j/2$ (in every cohort) while the aggregate number of workers investing in education is $N(0) = (\gamma_1 + \gamma_2)/2$. Define $\xi(t)$ as the fraction of the age-t skilled job-seekers who are from community 1:

$$\xi(t) = \frac{N_1(t)}{N(t)}, \qquad \text{and} \qquad 1 - \xi(t) = \frac{N_2(t)}{N(t)}. \tag{A2}$$

We want to show that $\xi(t)$ is uniquely identified. Using (10), along with (A1), yields:

$$\frac{\dot{\xi}(t)}{\xi(t)} + \varphi \frac{\dot{\xi}(t)}{1 - \xi(t)} = (1 - \varphi) \frac{K}{T - t}.$$

Removing the time-derivatives (using that $\dot{\xi}/\xi = d \ln \xi / dt$ and $\dot{\xi}/(1 - \xi) = -d \ln (1 - \xi)/dt$) and integrating yields

$$\frac{\xi(t)}{[1 - \xi(t)]^{\varphi}} = \left(\frac{T}{T - t}\right)^{K(1-\varphi)} \mathcal{C}(\gamma_1, \gamma_2) \tag{A3}$$

where $\mathcal{C}(\cdot, \cdot)$ is a strictly positive continuous function. Inspection of equation (A3) reveals that it has a unique solution $\xi(t) \in (0, 1)$ for every $t < T$ which varies continuously with γ_1 and γ_2. Note then that $F_j(t) = 1 - N_j(t)/N_j(0)$. Substituting for $N_j(t)$ and $N_j(0)$ then gives that

$$F_1(t) = 1 - \xi(t) \frac{(\gamma_1 + \gamma_2)}{\gamma_1} \left(\frac{T - t}{T}\right)^K,$$

while a similar expression holds for $j = 2$. To see the second part, recall that $a_j(\cdot)$ is the hazard rate associated with $F_j(\cdot)$; hence, $\ln(1 - F_j(t)) = -\int_0^t a_j(s)\, ds$. The result then follows from (10). Finally, if $\gamma_j = 0$, then by Assumption 1, $a_k(\cdot) = a(\cdot)$, which immediately implies $F_k(\cdot) = F(\cdot)$. ■

Proof of Proposition 3. Note that $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ are the same function. Hence any fixed-point, γ^*, for $\gamma_j(\cdot)$ is a symmetric equilibrium. But $\gamma_j(\cdot)$ maps the unit interval into itself and is continuous. Hence such a fixed-point exists. Moreover, as shown after Lemma (5), it follows that $\gamma_j'(\cdot) < 0$ at any symmetric point, which implies uniqueness. The waiting time $\bar{t}_0$ follows since $\gamma_1 = \gamma_2$ implies that $F_1(\cdot) = F_2(\cdot) = F(\cdot)$, defined in (7), which has the expected waiting time $T/(1 + K)$. Independence of v can be seen by noting that γ^* satisfies $G^{-1}(\gamma^*)\pi(\gamma^*) = \Delta w \left(T - \bar{t}_0\right)$ which does not depend on v. ■

Proof of Lemma 4. Note that as an adding-up identity,

$$\sum_{j=1,2} \frac{\gamma_j}{2} F_j(t) = F(t) \sum_{j=1,2} \frac{\gamma_j}{2}, \quad \text{for all } t \in [0, T], \tag{A4}$$

since both sides measure the number of skilled workers who have found good jobs by age t (recall that $\gamma_j/2$ is the number of workers from community j who invest in education). Using that $[1 - F_1(t)] = [1 - F_2(t)]^{\varphi}$ (Lemma 2) to substitute for F_1 in (A4), and the differentiating w.r.t. γ_1 (noting that $\varphi = (1 - \upsilon) + \upsilon \frac{\gamma_1}{\gamma_2}$), yields

$$\frac{\partial F_2(t)}{\partial \gamma_1} = \frac{F(t) - F_1(t) + (1 - F_2(t))^{\varphi} \upsilon \frac{\gamma_1}{\gamma_2} \ln(1 - F_2(t))}{\varphi \gamma_1 (1 - F_2(t))^{\varphi - 1} + \gamma_2}.$$

If $\gamma_1 \geq \gamma_2$, then $F_1(t) \geq F(t)$ whereby the above derivative is *strictly negative*. On the other hand, when $\gamma_1 < \gamma_2$ the sign is ambiguous. Similarly, differentiating w.r.t. γ_2 to obtain the own-effect yields

$$\frac{\partial F_2(t)}{\partial \gamma_2} = \frac{F(t) - F_2(t) - (1 - F_2(t))^{\varphi} \upsilon \left(\frac{\gamma_1}{\gamma_2}\right)^2 \ln(1 - F_2(t))}{\gamma_1 (1 - F_2(t))^{\varphi - 1} \varphi + \gamma_2}.$$

If $\gamma_1 \geq \gamma_2$, then $F(t) \geq F_2(t)$, whereby the derivative is *strictly positive*. On the other hand, when $\gamma_1 < \gamma_2$ the own effect is ambiguous. ■

Proof of Lemma 5. By inspection of Figure 1 it is clear that local stability (of any equilibrium) corresponds to the condition $\partial \gamma_j / \partial \gamma_k > \left(\partial \gamma_k / \partial \gamma_j\right)^{-1}$. Totally differentiating (14), which defines $\gamma_j(\gamma_k)$, gives

$$\frac{\partial \gamma_j}{\partial \gamma_k} \left[\frac{\pi(\gamma_j)}{g(\theta_j)} + \Delta w \frac{\partial \bar{t}_j}{\partial \gamma_j} + \pi'(\gamma_j) \theta_j \right] = -\Delta w \frac{\partial \bar{t}_j}{\partial \gamma_k}. \tag{A5}$$

The parenthesis is positive by (15), and, by Lemma 4, $\partial \bar{t}_j / \partial \gamma_k > 0$ at a symmetric equilibrium. At a symmetric equilibrium, $\partial \gamma_j / \partial \gamma_k = \partial \gamma_k / \partial \gamma_j < 0$, and the stability it can be stated as $\partial \gamma_j / \partial \gamma_k > -1$, or $\left|\partial \gamma_j / \partial \gamma_k\right| < 1$. ■

Proof of Proposition 6. Using that the aggregate output of a cohort can be written in the linear form given in the text, an efficient pair (γ_1, γ_2) can be characterized as a solution to

$$\max_{\gamma_1, \gamma_2} \left\{ T w_L + \sum_{j=1,2} \frac{\gamma_j}{2} \left[(T - \bar{t}_0) \Delta w \right] - \sum_{j=1,2} \frac{C(\gamma_j)}{2} \right\}$$

where $C(\cdot)$ is defined in (17). Since the objective function can be written in this separable form, is possible to optimize w.r.t. γ_1 and γ_2 separately, leading to the same set of maximizers. ■

Proof of Proposition 7. The analysis is facilitated by having explicit expressions for the derivatives studied in Lemma 4. It turns out that the cross- and the own effects differ only by sign: *at a symmetric equilibrium*

$$\frac{\partial \bar{t}_j}{\partial \gamma_k} = -\frac{\partial \bar{t}_j}{\partial \gamma_j} = \frac{vT}{2\gamma^*}.$$

(A proof is available on request.) Using this, along with Assumption 3, to simplify $\left|\partial \gamma_j / \partial \gamma_k\right|$ given in (16) yields the following expression for the slope, which we henceforth denote Σ,

$$\Sigma \equiv \left|\frac{\partial \gamma_j}{\partial \gamma_k}\right| = \frac{v\,(1+K)\,/K}{2\,(1 - g\theta^* a) - v\,(1+K)\,/K}. \tag{A6}$$

where θ^* is the cut-off education cost (in both communities), which is positively related to γ^* through the identity $\gamma^* = G(\theta^*)$. From this we see that a generic parameter α has two effects on Σ. One is through θ^*; this is the "scale effect". The second is the direct effect $\partial\Sigma/\partial\alpha$ (holding θ^* constant); this is the "sensitivity effect".

Note that an increase in scale is destabilizing, $\partial\Sigma/\partial\theta^* > 0$. Consider therefore the impact on scale of the symmetric equilibrium. Using that $\bar{t}_0 = T/\,(1+K)$ and that $\pi\,(\gamma) = e^{-a\gamma}$, θ^* can be characterized as follows as $\theta^* e^{-ag\theta^*} = \Delta w T K/\,(1+K)$. Taking logs and totally differentiating immediately yields that

$$\frac{\partial \theta^*}{\partial g} = \delta_0 a\theta > 0, \quad \frac{\partial \theta^*}{\partial \Delta w} = \frac{\delta_0}{\Delta w} > 0, \quad \frac{\partial \theta^*}{\partial K} = \frac{\delta_0}{K\,(1+K)} > 0, \quad \frac{\partial \theta^*}{\partial a} = \delta_0 g\theta > 0,$$

where $\delta_0 \equiv \theta^*/\,(1 - ag\theta^*)$ is strictly positive by (15). This proves part 1. Turning to the sensitivity effects, $\partial\Sigma/\partial\alpha$, inspection of (A6) immediately reveals that Σ increases in g, a and v, but decreases in K. This proves part 2. ■

Proof of Proposition 8. Treating (14) for $j, k = 1, 2$ now as an equation system, and using that (A5) provides an expression for $\gamma_j'\,(\gamma_k)$, (19) is obtained by straightforward comparative statics. In order to derive the direct effects, $\partial\gamma_j/\partial\alpha$, it is useful to rewrite (14) using Assumption 3 as $\gamma_j e^{-a\gamma_j} = g\Delta w\left(T - \bar{t}_j\left(\gamma_j, \gamma_k\right)\right)$. Note that K and v enter only through the waiting time $\bar{t}_j$. (The other parameters on the other hand do not directly affect the waiting times.) Taking logs and totally differentiating this equation then yields that

$$\frac{\partial \gamma_j}{\partial g} = \frac{\delta_1}{g} > 0, \quad \frac{\partial \gamma_j}{\partial \Delta w} = \frac{\delta_1}{\Delta w} > 0, \quad \frac{\partial \gamma_j}{\partial K} = -\frac{\delta_1}{\left(T - \bar{t}_j\right)}\frac{\partial \bar{t}_j}{\partial K},$$

$$\frac{\partial \gamma_j}{\partial a} = \delta_1 \gamma_j > 0, \quad \text{and} \quad \frac{\partial \gamma_j}{\partial v} = -\frac{\delta_1}{\left(T - \bar{t}_j\right)}\frac{\partial \bar{t}_j}{\partial v},$$

where

$$\delta_1 \equiv \frac{1}{\gamma_j} - a + \frac{1}{(T - \bar{t}_j)} \frac{\partial \bar{t}_j}{\partial \gamma_j}$$

is strictly positive by (15). To determine the signs of $\partial\gamma_j/\partial K$ and $\partial\gamma_j/\partial v$ we need to determine the impacts on the waiting times. Applying the same technique as in the proof of Lemma 4, using (7) and (10), reveals that $\partial\bar{t}_j/\partial K < 0$ for $j = 1, 2$, while $\partial\bar{t}_j/\partial v$ is strictly negative for $j = 1$ and strictly positive for $j = 2$ when $\gamma_1 > \gamma_2$. ■

References

Acemoglu, D. & Pischke, J. S. (2001), 'Changes in wage structure, family income, and children's education', *European Economic Review* **45**, 890–904.

Becker, G. & Tomes, N. (1986), 'Human capital and the rise and fall of families', *Journal of Labor Economics* **107**, 123–150.

Bénabou, R. (1993), 'Workings of a city: Location, education, and production', *Quarterly Journal of Economics* **108**, 619–652.

Bénabou, R. (1996), 'Equity and efficiency in human capital investment: The local connection', *Review of Economic Studies* **63**, 237–264.

Björklund, A. & Jäntti, M. (1997), 'Intergenerational income mobility in Sweden compared to the United States', *American Economic Review* **87**, 1009–1018.

Björklund, A. & Jäntti, M. (2000), 'Intergenerational mobility of socio-economic status in comparative perspective', *Nordic Journal of Political Economy* **22**, 3–32.

Blanden, J., Goodman, A., Gregg, P. & Machin, S. (2002), 'Changes in intergenerational mobility in Britain'. Mimeo, University College London.

Blanden, J., Gregg, P. & Machin, S. (2002), 'Education and family income'. Mimeo, University College London.

Bulow, J. I. & Summers, L. H. (1986), 'A theory of dual labor markets with application to industrial policy, discrimination, and keynesian unemployment', *Journal of Labor Economics* **4**, 376–414.

Calvó-Armengol, A. & Zenou, Y. (2001), 'Job matching, social network and word-of-mouth communication'. CEPR Discussion Paper Series 2797.

Cameron, S. & Heckman, J. J. (1998), 'Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males', *Journal of Political Economy* **106**, 262–333.

Cameron, S. & Heckman, J. J. (2001), 'The dynamics of educational attainment for black, hispanic, and white males', *Journal of Political Economy* **109**, 455–499.

Carneiro, P. & Heckman, J. J. (2002), 'The evidence of credit constraints in post-secondary schooling', *Economic Journal* **112**, 705–734.

Checchi, D., Ichino, A. & Rustichini, A. (1999), 'More equal but less mobile? Education financing and intergenerational mobiity in Italy and the US', *Journal of Public Economics* **74**, 351–393.

Corcoran, M., Gordon, R., Laren, D. & Solon, G. (1989), 'Effects of family and community background on men's economic status'. Working Paper No. 2896, National Bureau of Economic Research.

Dearden, L., Machin, S. & Reed, H. (1997), 'Intergenerational mobility in Britain', *Economic Journal* **107**, 47–64.

Durlauf, S. N. (1996), 'A theory of persistent income inequality', *Journal of Economic Growth* **1**, 75–93.

Ellwood, D. T. & Kane, T. J. (2000), Who is getting into college? Family background and the growing gaps in enrollment, *in* S. Danziger & J. Waldfogel, eds, 'Securing the Future', Russell Sage, New York.

Fertig, A. (2001), 'Trends in intergenerational earnings mobility'. Center for Research on Child Well-Being, Working Paper No. 01-23.

Granovetter, M. (1995), *Getting a Job*, 2 edn, The University of Chicago Press, Chicago.

Gregg, P. & Wadsworth, J. (1996), 'How effective are state employment agencies? Jobcentre use and job matching in Britain', *Oxford Bulletin of Economics and Statistics* **58**, 43–67.

Holzer, H. J. (1988), 'Search method use by unemployed youth', *Journal of Labor Economics* **6**, 1–20.

Korpi, T. (2001), 'Good friends in bad times? Social networks and job search among the unemployed in Sweden', *ACTA Sociologica* **44**, 157–170.

Loury, G. C. (1981), 'Intergenerational transfers and the distribution of earnings', *Econometrica* **49**, 843–867.

Mayer, S. (1997), *What Money Can't Buy: Family Income and Children's Life Chances*, Harvard University Press.

Mayer, S. & Lopoo, L. (2001), 'Has the intergenerational transmission of economic status changed?'. Mimeo, University of Chicago.

Mulligan, C. B. (1997), *Parental Priorities and Economic Inequality*, The University of Chicago Press, Chicago.

Saint-Paul, G. (2001), 'The political economy of employment protection', *Journal of Political Economy* . Forthcoming.

Shapiro, C. & Stiglitz, J. (1984), 'Equilibrium unemployment as a worker discipline device', *American Economic Review* **74**, 433–444.

Solon, G. (1999), Intergenerational mobility in the labor market, *in* O. Ashenfelter & D. Card, eds, 'Handbook of Labor Economics', North-Holland, Amsterdam.

Wasmer, E. & Zenou, Y. (2001), 'Equilibrium search unemployment with explicit spatial frictions'. The Research Institute of Industrial Economics, Stockholm, Mimeo.

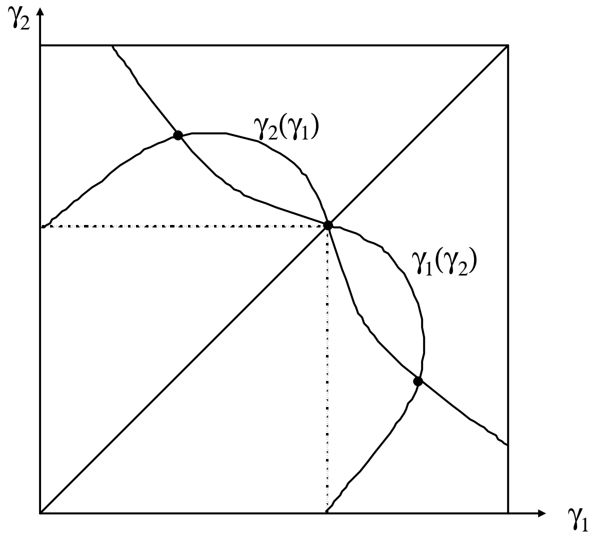

Figure 1:

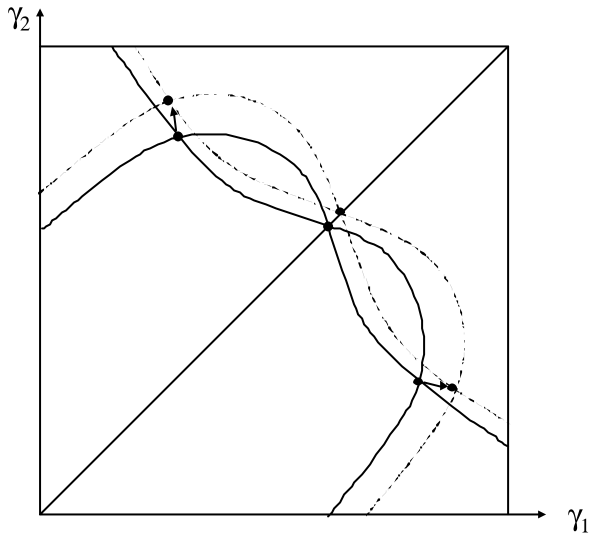

Figure 2: