

NBER WORKING PAPER SERIES

EDUCATION, INCOME
DISTRIBUTION, AND GROWTH:
THE LOCAL CONNECTION

Roland Bénabou

Working Paper No. 4798

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1994

I am grateful to Olivier Blanchard, Patrick Bolton, Peter Diamond, Dick Eckaus, Jonathan Gruber, Michael Kremer, Glenn Loury, Julio Rotemberg, Richard Startz and Jean Tirole for useful comments. Financial support from the NSF is gratefully acknowledged. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

EDUCATION, INCOME
DISTRIBUTION, AND GROWTH:
THE LOCAL CONNECTION

ABSTRACT

This paper develops a simple model of human capital accumulation and community formation by heterogeneous families, which provides an integrated framework for analyzing the local determinants of inequality and growth. Five main conclusions emerge. First, minor differences in education technologies, preferences, or wealth can lead to a high degree of stratification. Imperfect capital markets are not necessary, but will compound these other sources. Second, stratification makes inequality in education and income more persistent across generations. Whether or not the same is true of inequality in total wealth depends on the ability of the rich to appropriate the rents created by their secession. Third, the polarization of urban areas resulting from individual residential decisions can be quite inefficient, both from the point of view of aggregate growth and in the Pareto sense, especially in the long run. Fourth, when state-wide equalization of school expenditures is insufficient to reduce stratification, it may improve educational achievement in poor communities much less than it lowers it in richer communities; thus average academic performance and income growth both fall. Yet it may still be possible for education policy to improve both equity and efficiency. Fifth, because of the cumulative nature of the stratification process, it is likely to be much harder to reverse once it has run its course than to arrest it at an early stage.

Roland Bénabou
Department of Economics
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, MA 02139
and NBER

Introduction

The accumulation of human capital underlies the evolution of both income inequality and productivity growth. As demonstrated most vividly by the physical blight and social pathology of inner-city schools, certain fundamental inputs in this process are of a local nature. They are determined neither at the level of individual families nor that of the whole economy, but at the intermediate level of communities, neighborhoods, firms or social networks. Not only is this the case with school resources when funding is decentralized, but also with many forms of "social capital": peer effects, role models, job contacts, norms of behavior, crime, and so on. Through these fiscal and sociological spillovers, a child's education is determined in large part by the set of adults to which she is "connected". Therefore, the next generation's distribution of skills and incomes is directly shaped by the manner in which today's population organizes itself into differentiated clusters.

The aim of this paper is twofold. First, to provide a unified analysis of the forces that tie together socioeconomic segregation, income distribution, and productivity growth. We develop a simple model where heterogeneous families form communities, choose local public expenditures, and accumulate human capital. It allows us to elucidate the roles played by endowments, preferences, the technology of education, capital markets, school funding, neighborhood effects, and discrimination. Its synthetic quality also enables us to incorporate some of the main insights from the previous literature, while obtaining several new ones. The second objective of this paper is to further our understanding of two important questions. One is the effect of stratification on the transmission of inequality across generations; it bears directly on the problem of inner-city poverty, and on social mobility in general. The other central issue concerns education finance policy: we examine some of the distributional and aggregate implications of the growing trend towards court-mandated, state-wide equalization of school expenditures.

The literature to which the paper belongs has its sources in the classic works of Tiebout (1956) on local public goods and Schelling (1978) on segregation and externalities. More recent prominent influences include the work of Loury (1977), (1987) on racial inequality and of sociologists such as Coleman (1988), Wilson (1987) and Jencks and Mayer (1990) on group interactions. Another important motivation is the current debate over local disparities in school funding, perhaps best publicized by Kozol (1991). The model developed here builds on De Bartolome (1990), Bénabou (1993), (1992) and Durlauf (1992). But the paper

is also closely related to Borjas (1992a), (1992b), Durlauf (1993), Tamura (1993) and Lundberg and Startz (1993), through the link between community composition and the persistence of inequality; and to Glomm and Ravikumar (1992), Fernandez and Rogerson (1992), (1993b) and Cooper (1992), through the issue of education finance reform.

We now outline the structure of the paper and its main results. The model is presented in Section 1. In Section 2, we identify the forces which promote or hinder segregation and explain their interactions. We show how even very small differences in education technologies, preferences or wealth lead to a high degree of stratification. Capital market imperfections are not necessary, but even minor ones will compound the other factors. The rest of the paper studies the potentially large effects of these small causes. Focusing first on productivity, we examine in Section 3 how the total surplus generated by a metropolitan area reflects its organization into local communities. In particular, we explain why the typical pattern of city-suburb polarization can be very inefficient.

We then take up the issue of whether socio-economic segregation makes inequality more persistent. We demonstrate in Section 4 how income convergence is slowed down, how ghettos arise, and how in the long run this may reduce income growth for all families. We present in particular a very simple prototype of the "local poverty trap" and "self-defeating secession" phenomena, which makes apparent the general principles at work in more complex models (e.g., Bénabou (1993), Durlauf (1992), (1993)). But because our framework allows for financial bequests as well as human capital, we are also able to point out an important qualification which the previous literature has generally overlooked. While stratification exacerbates inequality in education and income, the same need not be true for inequality in total wealth. The determinant factor is the cost paid by the rich to separate themselves from the poor, or conversely the extent to which they are able to appropriate the rents generated in the process of segregation. We discuss various collective practices and institutions by which this may come about, showing in particular how *de jure* racial segregation turns into *de facto* (economic) segregation once legal barriers to mobility are lifted.

The other main issue studied in the paper is education finance. We show in Section 5 that the decentralization of school funding and control constitutes an additional segregating force, and that in general it need not improve efficiency. In Section 6 we then examine the implications of policies, adopted by a growing number of states, aimed at reducing disparities in education expenditures between rich and poor

communities. Bénabou (1992) and Fernandez and Rogerson (1993b) predict very significant gains to moving from a system where school funding is based on local resources to a national or state-wide scheme. In addition to reducing inequality, this would raise the economy's long-term output, or even growth rate; given a low enough intergenerational discount rate, it would be Pareto-improving. In this paper we temper this optimistic scenario by downplaying capital market imperfections, and emphasize instead the interaction between purchased inputs and social spillovers in education. Indeed, several pieces of evidence seem to cast doubt on the effectiveness of redistributive funding in raising the performance of poor schools, hence in reducing earnings inequality and augmenting aggregate surplus. We show how a simple version of our model can help explain some of these puzzles. We determine when equalization works and when it is counterproductive, and, in the latter case, identify alternative policies which may still improve both equity and efficiency. Because these policies work through changes in community composition, however, they may be constrained by a form of irreversibility which we show to be inherent in the stratification process.

We conclude this preamble with some suggestive evidence on metropolitan stratification and growth. While the effects of neighborhood and school composition on individual outcomes have been extensively documented, there is as yet no empirical study of their aggregate implications. Rusk (1993) presents some provocative data which bears on the issue; since it is very incomplete, we include it as general motivation for the paper rather than supporting evidence for specific results. Table 1 examines fourteen metropolitan areas, selected by Rusk; the data is also plotted in Figure 1. It shows a strong positive relationship between the *lack of economic segregation*, measured by the ratio of central city to suburban mean incomes, and the *area's growth* in both per capita income (1969-1989) and total employment (1973-1988). The correlations are respectively .74 and .91, or .60 and .45, depending on whether we measure income disparities in 1989, as Rusk does, or in 1970, which may be more appropriate.¹ This striking stylized fact is no substitute for a systematic econometric analysis, and does not allow any inference about causality. But it does suggest that city-suburb stratification and metropolitan economic performance are interdependent processes, and that the underlying mechanisms deserve careful theoretical and empirical investigation.

¹I am grateful to Ed Glaeser for kindly providing the 1970 figures, from the Census Bureau's County and City Data Books. The above result seems opposite to that of Glaeser, Scheinkman and Shleifer (1993), who find that racial segregation in a city affects population growth positively. Note, however, that the focus is here on city-suburb economic dichotomy, rather than on racial segregation and population growth within city limits.

Metro Area	City/Suburb per capita income ratio (%)		Metro per capita income growth (%) 1969-1989	Metro employment growth (%) 1973-1988
	1970	1989		
Cleveland, Ohio	69	53	23	8
Detroit, Mich.	80	53	26	11
Columbus, Ohio	83	81	34	40
Milwaukee, Wis.	84	62	25	22
Harrisburg, Pa.	87	72	33	29
Louisville, Ky.	91	79	30	21
Richmond, Va.	91	83	45	46
Nashville, Tenn.	92	98	49	55
Syracuse, N.Y.	98	77	29	27
Madison, Wis.	104	95	34	51
Houston, Tex.	107	89	36	67
Indianapolis, Ind.	108	90	32	34
Raleigh, N.C.	119	103	62	81
Albuquerque, N. Mex.	148	118	41	75

Table 1: City-Suburb Inequality and Metropolitan Area Performance.

Source: see Rusk (1993), and footnote 1

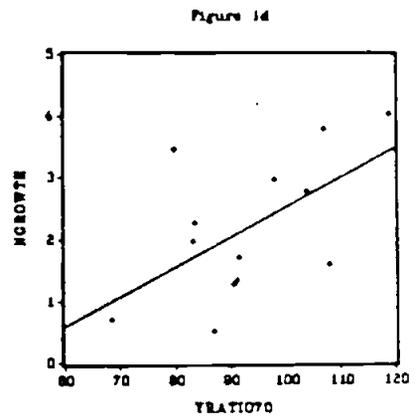
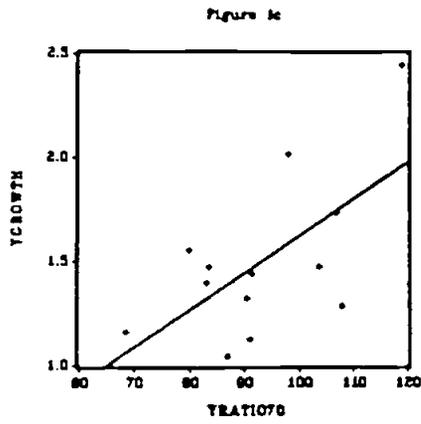
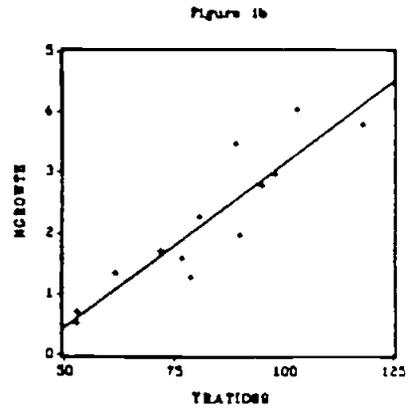
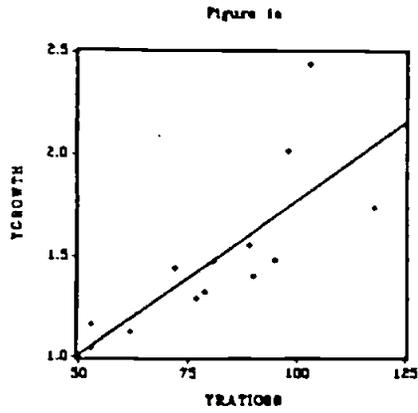


Figure 1: City-Suburb Inequality and Metropolitan Area Performance
 YRATIO= relative city/suburb mean income; YGROWTH=annual growth rate of metropolitan income per capita; NGROWTH= annual growth rate of metropolitan employment. All values are percentages

1 The model

1.1 Agents and communities

There is a continuum of families, with unit measure. Parents are of two types, A and B (rich and poor. White and Black, etc.), with human capital endowments $h_A > h_B$. The proportions of the two types are n and $1 - n$. The average level of human capital is denoted $\bar{h} = n h_A + (1 - n) h_B$. These agents live in a city composed of two towns or communities, $j = 1, 2$, each holding the same number $1/2$ of single-family homes. The inelasticity of land supply and population density simplifies the analysis but is not essential. All land belongs to absentee landowners; departures from this neutral allocation are considered later on. The proportion of rich, or type A agents in each community is denoted as x^j , and community 1 is defined as the richer one: $x^1 \geq 2n - x^1 = x^2$. In the American context the two locations would be referred to as “the suburbs” and “the urban center”; in Europe it would be the reverse.

A model with two types and two communities is the minimal framework in which to study how agents associate in the provision of formal and informal education. It can also be viewed as a representative “slice” of a city with many types and communities, where we focus on any pair of contiguous towns and their populations; almost all results extend to such a model. Thus we shall occasionally make the convenient assumption that $\Delta h/\bar{h} \equiv (h_A - h_B)/\bar{h}$ is small.

1.2 Preferences and technologies

There are two periods. A parent with type $h \in \{h_A, h_B\}$ initially chooses a community C^j , $j = 1, 2$, so as to maximize the resulting utility $U^j(h)$:

$$\begin{aligned} U^j(h) &\equiv \max_d U(c, c', h'), \text{ subject to} \\ c + \rho^j + t^j(h) &= \omega(h) + d \\ c' + P(h, d) &= y(h) \\ h' &= F(h, L^j, E^j) \end{aligned} \tag{1}$$

In the first period, she consumes c and pays rent ρ^j plus taxes $t^j(h)$ out of her initial resources $\omega(h)$, augmented by her chosen level of debt d . The interest rate $r(h, d)$ may depend on her type as well as the amount borrowed or saved. In the second period, the resources c' available for consumption or bequest

(both interpretations are possible) equal second period income $y(h)$, minus debt repayments $P(h, d) \equiv d(1 + r(h, d))$. Finally, the child's human capital h' is determined by that of the parent (through at-home education), by the quality L^j of social interactions in the chosen community, and by the resources devoted to its schools, measured by the per student budget E^j .²

1.3 Neighborhood effects

The local spillover L^j captures the non-fiscal channels through which a child's acquisition of skills is affected by the social mix of neighboring families. These sources of "social capital" (Loury (1977), Coleman (1988)) include: peer effects between students from different backgrounds, both in and out of the classroom (Banerjee and Besley (1991)); the fact that neighboring adults enforce norms of behavior and provide role models, as well as networking contacts, for the young (Wilson (1987), Streufert (1991), Montgomery (1991)); and crime or other activities which disrupt education (Sah (1991)).³ Without loss of generality, we assume that $L^j = L(x^j; h_A, h_B)$ is homogeneous of degree one in the distribution of human capital, and that if all adults in C^j share the same level h , then $L^j = h$. We also require L to increase with any improvement (in the sense of first-order stochastic dominance) in the local distribution: $L'(x) \geq 0$, and L rises with h_A and h_B . One can thus think of L^j as an *average* of local residents' levels of human capital; its value for a representative sample of population will be denoted $\bar{L} \equiv L(n; h_A, h_B)$.

Another key feature of the spillover is its response to a *mean-preserving spread* in the distribution. How costly, or how beneficial is heterogeneity within a school or neighborhood? Put in another way, do stronger individuals tend to "pull up" the average to their level, or do weaker ones tend to "drag it down" to theirs? Clearly, it is important to allow both cases.⁴ A convenient specification which we shall sometimes use is the CES index $L(x) = (x h_A^{\frac{\epsilon-1}{\epsilon}} + (1-x) h_B^{\frac{\epsilon-1}{\epsilon}})^{\frac{\epsilon}{\epsilon-1}}$. When $1/\epsilon > 0$, individual levels of human capital are complements, L is convex in x , and heterogeneity is a source of loss: $L < \bar{h}$. When $1/\epsilon < 0$, individual

²More generally, $F(h, L, E)$ should be understood as the child's earnings potential, net of her optimally chosen studying effort (measured in the numeraire good). Incorporating private purchases of education is also straightforward: c and E are replaced in (1) by $c + e$ and $E + e$, and parents maximize over both e and d .

³Jencks and Mayer (1990) provide an extensive survey of the empirical evidence on neighborhood effects, and Manski (1993) a discussion of methodology. Recent studies include Datcher (1982), Dynarski et al. (1989), Corcoran et al. (1990), Crane (1991), Borjas (1992a), (1992b), Case and Katz (1991), and Brooks-Gunn et al. (1993).

⁴Dynarski, Schwab and Zampelli (1989) find that average student performance in a school district rises, *ceteris paribus*, with dispersion among family incomes. Hamilton and Macauley (1991), on the other hand, find that income dispersion raises the unit cost of education. Brooks-Gunn et al. (1993) find that it is primarily the upper tail of a neighborhood's income distribution which affects the development of children and adolescents raised there.

levels are substitutes, L is concave in x , and heterogeneity is a source of gains: $L > \bar{h}$. The Leontieff (minimum) and “best-shot” (maximum) cases are obtained in the limit as $1/\varepsilon$ goes to $+\infty$ or $-\infty$. More generally, the concavity of $L(x)$ will be an important determinant of the efficiency of equilibrium.⁵

1.4 Local school funding

The other local input into education is decentralized school expenditures.⁶ Given the distribution $(x^j, 1-x^j)$ of rich and poor agents in the community –hence, given social capital L^j and rents ρ^j – their preferences over school budgets and taxation are aggregated by some local *political mechanism* into funding decisions $E^j = E(x^j)$ and tax schedules $t^j(h, \rho) = t(h, \rho, x^j)$, subject to the budget constraint $x^j t(h_A, \rho^j, x^j) + (1-x^j) t(h_B, \rho^j, x^j) = E(x^j)$. We thus allow taxes to depend on individual income and property values.⁷ The political mechanism in question could be majority voting, the outcome of lobbying, or delegation to an efficient local planner. For the moment we shall leave it unspecified, so as to identify the general features of mechanisms which promote or hinder segregation.

1.5 Equilibrium

Let us rewrite the utility of an adult of type h , living in a community with percentage x of rich households and paying ρ in land rent, as:

$$V(h, \rho, x) = \max_d \{U(\omega(h) + d - \rho - t(h, \rho, x), y(h) - P(h, d), F(h, L(x), E(x)))\} \quad (2)$$

Equilibrium in the land market will result in stratification if the rich (in human capital) are willing or able to bid more than the poor for land in a richer community. Formally, this simple *sorting condition* in the space of community quality and price means that a rich family’s iso-utility curve, or bid-rent, is steeper

⁵For a more detailed discussion, see Bénabou (1992). Similar functional forms are used by Cornes (1993) in studying the provision of public goods with non-additive individual contributions.

⁶As for social composition, the link between expenditures and students’ achievement or future earnings is the subject of much empirical debate. See Hanushek (1986) and Card and Krueger (1992) for opposing views.

⁷Note that the E^j and $t^j(h, \rho)$ chosen by C^j residents may depend on the *equilibrium* rent ρ^j (as opposed to an individual’s bid ρ), through its effects on their wealth and on potential revenues from alternative forms of taxation. To keep the notation simple, this dependence on j is subsumed in the dependence on x^j .

than a poor one's:

$$R_x(h, \rho, x) \equiv \left. \frac{d\rho'}{dx'} \right|_{V(h, \rho', x')=V(h, \rho, x)} = \frac{V_x(h, \rho, x)}{-V_\rho(h, \rho, x)} \text{ increases with } h \quad (3)$$

In that case, the slightest divergence from the symmetric allocation $x^1 = x^2 = n$ (which is always an equilibrium) in favor of community 1 will set in motion a cumulative process: rich agents outbid poor ones for land in C^1 , thereby further raising x^1, ρ^1 and lowering x^2, ρ^2 . This leads to more displacements of the poor by the rich in C^1 and more concentration of poverty in C^2 , until at least one community is completely homogeneous. The three possible configurations are indicated on Figure 1, together with the equilibrium conditions on ρ^1 and ρ^2 .⁸

Proposition 1 1. If $R_{hx}(h, \rho, x) > 0$ for all ρ, x , the unique stable equilibrium is stratified. If $n \leq 1/2$, the rich all live in community 1 ($x^1 = 2n, x^2 = 0$); if $n \geq 1/2$, the poor all live in community 2 ($x^1 = 1, x^2 = 2n - 1$). The symmetric equilibrium $x^1 = x^2 = n$ is unstable.

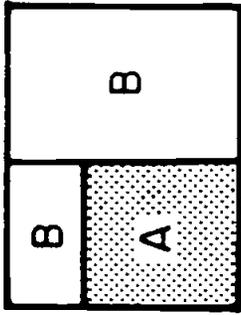
2. If $R_{hx}(h, \rho, x) < 0$ for all ρ, x , the unique equilibrium is completely integrated (i.e. symmetric), and it is stable.

The proof is given in appendix. Examining the marginal rate of substitution between first-period consumption and child education thus brings to light the economic forces which lead to stratification or integration. Using the Euler equation for the optimal level of borrowing, we have:

$$R_x(h, \rho, x) = \left(\frac{U_3(c, c', h')}{U_2(c, c', h')} \cdot \frac{F_L \cdot L'(x) + F_E \cdot E'(x)}{P_d(h, d)} - t_x(h, \rho, x) \right) (1 + t_\rho(h, \rho, x))^{-1} \quad (4)$$

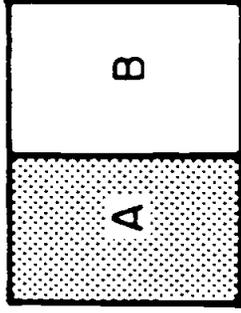
Equation (4) incorporates the contributions of technology, preferences, capital markets and local public goods. The corresponding components will be discussed in turn.

⁸As in much of the literature on community composition (e.g., De Bartolome (1990), Wheaton (1993)) we only determine the rent differential which ensures that no agent wants to switch communities. The base level of these rents could be pinned down by relaxing one of the simplifying assumptions of inelastic land supply and fixed number of agents of each type, but this would needlessly complicate the model.



Case $n < 1/2$

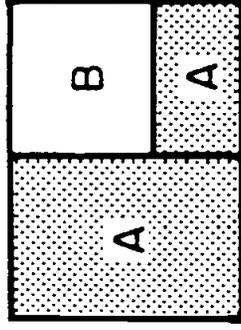
$$V(h_B, \rho^1, x^1) = V(h_B, \rho^2, x^2)$$



Case $n = 1/2$

$$V(h_B, \rho^1, x^1) - V(h_B, \rho^2, x^2) \leq 0$$

$$V(h_A, \rho^1, x^1) - V(h_A, \rho^2, x^2) \geq 0$$



Case $n > 1/2$

$$V(h_A, \rho^1, x^1) = V(h_A, \rho^2, x^2)$$

FIGURE 2

2 The determinants of stratification

2.1 Local complementarities

The simplest version of the model is that where: (a) the only local input into education is community quality, $F = F(h, L)$, so that no collective decisions are involved;⁹ (b) capital markets are perfect, $P(h, d) = d(1+r)$; (c) utility is linear, $U(c, c', h') = c + (c' + h') / (1+r)$, or more generally c' and h' enter additively, so that U_3/U_2 is constant. This last assumption seems especially appropriate when c' is a bequest. The bid-rent slope now simplifies to:

$$R_x(h, \rho, x) = \frac{F_L(h, L)}{1+r} L'(x) \quad (5)$$

Stratification occurs if families with higher human capital are more sensitive to neighborhood quality than those with lower human wealth. In the words of Baumol (1967):

“The individual’s remedy intensifies the community’s problems and each feeds upon the other. Those who leave the city are usually the very persons who care and can afford to care—the ones who maintain the houses, who do not commit crimes, and who are most capable of providing the taxes needed to arrest the process of urban decay. Their exodus therefore leads to further deterioration in urban conditions and so induces yet another wave of emigration, and so on.”

Proposition 2 *A small amount of complementarity between family human capital and community quality ($F_{hL} > 0$) is sufficient to cause stratification. The distribution of financial wealth is irrelevant.*

This simplest specification of preferences and technologies will be referred to as the “basic model”. In contrast to most of the literature on inequality, it shows that there might be very *little role for redistributive policies* to affect either the distribution of educational attainment or the efficiency of equilibrium.

⁹This model also admits two interesting alternative interpretations. One is that agents care about community quality L for reasons other than its impact on their child’s education; crime is a good example, with richer agents caring most about it. The second is that $F = F(h, E)$ and communities finance schools by taxing income at a fixed rate (such is then case when preferences are logarithmic; Glomm and Ravikumar (1992)). Then $E(x) = r^* L(x)$, where $L(x) \equiv x \omega(h_A) + (1-x) \omega(h_B)$ is simply average income in period 1.

2.2 Capital market imperfections

In reality, it may not be easy for a poor family who values education highly to borrow enough to move into a rich, well-educated community. We now show that differences in “ability” to pay for community quality tend to complement, or even replace, differences in “willingness” to pay in generating economic segregation.¹⁰ We abstract from inter-family differences in tastes or returns to education, $U_3/U_2 = 1$, $F = F(L)$, and allow instead frictions in credit markets: $R_x(h, d) = F'(L) L'(x) / P_d(h, d)$.

Proposition 3 *Small imperfections in capital markets, resulting in a higher opportunity cost of funds for poor families ($P_{hd} + d'(h) P_{dd} < 0$) are sufficient to cause stratification.*

The result arises most simply when each family faces a constant opportunity cost of funds, with $r_A < r_B$ in (5). This could reflect different monitoring technologies, or a tax subsidy to home ownership relative to renting (interest deductions), when only A families can afford a downpayment.

A better way is to endogenize this differential. Let everyone face the *same* interest schedule $P(d) = d(1 + r(d))$, with increasing marginal cost of borrowing (or decreasing marginal return on savings) $P'' > 0$, due to standard asymmetric information problems. Let agents work only in their first period of life, earning $\omega(h)$, with $\omega'(h) > 0$; during retirement, $y(h) = 0$. In order to live in the same community, less wealthy families must then borrow more than rich ones: d decreases with h for any given ρ , hence the result. The simplest case is that of a wedge between borrowing and lending rates, as in Galor and Zeira (1993); conditions on endowments which ensure that A agents are always lenders and B agents always borrowers are easily found. Alternatively, the imperfection could take the form of a borrowing constraint, $d \leq \bar{d}$. When resources are such that it binds for the B 's but not for the A 's, the Euler condition is an inequality for the former but an equality for the latter, implying again $R_x(h_A, \rho, x) > R_x(h_B, \rho, x)$. But it bears repeating that, absent offsetting forces, segregation will occur *no matter how small* wealth differences and credit market imperfections.

¹⁰The empirical literature yields contradictory estimates of the sign of F_{hL} ; see Jencks and Mayer (1990). The most recent study (Brooks-Gunn et al. (1993)) finds $F_{hL} > 0$.

2.3 Wealth effects

We now focus on the term $(U_3/U_2)(c, c', h')$ in (4), which captures preferences for education relative to old age consumption or financial bequest. To isolate wealth effects, let $F = F(L)$ and capital markets be perfect. Finally, let parental utility be additively separable: $U(c, c', h') = u(c) + v(c') + w(h')$. Then:

$$R_{\mathbf{x}}(h, \rho, \mathbf{x}) = \frac{w'(F(L))}{v'(c')} F'(L) L' = \frac{w'(F(L))}{\tilde{u}'(z(h) - \rho)} F'(L) L'$$

where $\tilde{u}(z) \equiv \max_s \{u(z - s) + v(s(1 + r))\}$ and $z(h) \equiv \omega(h) + y(h)/(1 + r)$ is lifetime wealth.

Proposition 4 *Small differences in lifetime resources ($\omega'(h) + y'(h)/(1 + r) > 0$) are sufficient to cause stratification.*

When $F = F(h, L)$, the marginal rate of substitution is $U_3 / U_2 = w'(F(h, L)) / \tilde{u}'(z(h) - \rho)$. Stratification occurs if the reduction in the marginal utility of education due to better parental background h is less than the reduction in the marginal utility of consumption which results from higher financial wealth $z(h)$. This is the sense in which community composition L must be a *normal good*. Similar wealth effects underlie models where sorting occurs not through land rents –supply is infinitely elastic– but through different levels or rates of taxation across communities (e.g., Fernandez-Rogerson (1992)). Ceteris paribus, rich agents vote for high taxes to finance public goods, such as education, whereas poor ones prefer lower levels. This leads, through a Tiebout (1956) -like mechanism, to as much segregation by wealth as the number of communities will allow. Instead of rents or taxes, wages can also play the role of compensating differentials. Such is the case in Fershtman and Weiss's (1992) analysis of *social status*, where agents derive utility from being in a profession with a high proportion of well educated members. Wealthier agents are more willing to accept a lower remuneration of human capital, in exchange for higher status. Therefore more of them become highly educated, and the labor market stratifies. We show in appendix how variations of our model yield simple versions of Fernandez-Rogerson (1992) and Fershtman and Weiss (1992).

We have until now assumed that all land rents accrue to absentee landowners. This is a “neutrality” assumption with respect to the allocation of the capital gains and losses created by stratification in C^1 and C^2 . It also simplifies the problem, by making agents' initial wealth exogenous: $\omega(h) = \bar{\omega}(h)$. We now examine how relaxing this assumption affects the results. Let each A agent initially own ν_A units of land

in C^1 and η_A units in C^2 ; the corresponding endowments for a B agent are $\nu_B = (1/2 - n\nu_A)/(1 - n)$ and $\eta_B = (1/2 - n\eta_A)/(1 - n)$. Total wealth levels are $\omega(h_i) = \bar{\omega}(h_i) + \nu_i\rho^1 + \eta_i\rho^2$, $i = A, B$. If $\nu_A = \eta_A = 1/2$, both classes share equally in capital gains and losses; this is another neutral case where the results remain unchanged. The converse, perhaps more realistic case, is $\nu_A = 1/2n$, $\eta_A = 0$. More generally, consider any initial allocation where $\nu_A > \eta_A$, hence $\nu_B < \eta_B$. As communities become more concentrated, the wealth of the A 's rises while that of the B 's declines. Propositions 3 and 4 show that this *in itself* creates a segregating force, which can sustain stratification even when all others are absent.

2.4 Local public goods, political mechanisms, and discrimination

We defer until Section 5 specifying a particular mechanism through which communities choose taxes and school expenditures. But inspection of the corresponding terms in (4) already indicates a "natural" tendency toward stratification, with a presumption of efficiency gains, a la Tiebout (1956).

Proposition 5 *Political mechanisms whose outcome is such that expenditures $E(x)$ increase with x when $F_{hE} > 0$ (respectively, decrease when $F_{hE} < 0$) tend to produce stratification. So do mechanisms such that income tax rates fall with the proportion of rich agents in the community ($t_{hx}(h, \rho, x) < 0$) or with land values ($t_{h\rho}(h, \rho, x) < 0$).*

To the extent that a higher proportion of rich agents is reflected in taxation and expenditure decisions which are closer to their preferred choices, stratification will occur as agents "vote with their feet". Although not all public choice mechanisms are such that the rich become more powerful as they become more numerous (e.g., majority voting), this requirement seems both fairly weak and realistic.

An alternative interpretation of $t_{hx}(h, \rho, x) < 0$ or $t_{h\rho}(h, \rho, x) < 0$ is that the poor incur higher costs, relative to the rich, of joining a rich community. These costs, which are particularly relevant when the A 's and B 's correspond to different ethnic groups, can be pecuniary, due to discrimination by owners or intermediaries in the housing market, or non-pecuniary, such as harassment.¹¹ We come back to them when discussing inequality.

¹¹Of course, this component of t does not enter into budget constraints. Yet another interpretation is that of agents who have pure "tastes" for the ethnic mix of their community, as in Schelling (1978) or Miyao (1978).

3 Stratification and efficiency: a simple case

We now examine whether it is efficient for different classes of agents to segregate in education. The criterion we use is that of aggregate productivity or surplus: are the output gains of rich communities greater or smaller than the losses of poor ones? Distributional issues are taken up in Section 4.¹² To simplify the exposition, we abstract from school funding and taxes; they are incorporated in Section 5. The surplus generated by a community with a fraction x of rich households is thus $S(x)/2$, where:

$$S(x) \equiv xF(h_A, L(x)) + (1-x)F(h_B, L(x)) \quad (6)$$

A planner choosing population allocations so as to maximize the productivity of the whole metropolitan area would therefore maximize $S(x) + S(2n-x)$ over $0 \leq x \leq \min\{2n, 1\}$. If S is convex, the optimal solution is at a corner, and coincides with the stratified market equilibrium. On the other hand if S is concave, there are decreasing social returns to the concentration of human capital, and the optimum is interior and symmetric: $S'(x) = S'(2n-x)$, so $x = n$.¹³ The underlying intuition is apparent when comparing the gains $S(x^1) - S(n)$ from stratification in the rich community with the losses $S(n) - S(x^2)$ in the poor one. To assess the efficiency of equilibrium, we therefore evaluate S'' :

$$\begin{aligned} S' &= F(h_A, L) - F(h_B, L) + (x F_L(h_A, L) + (1-x) F_L(h_B, L)) L' ; \\ S'' &= 2(F_L(h_A, L) - F_L(h_B, L)) L' + (x F_{LL}(h_A, L) + (1-x) F_{LL}(h_B, L)) (L')^2 \\ &\quad + (x F_L(h_A, L) + (1-x) F_L(h_B, L)) L'' \end{aligned} \quad (7)$$

where $L = L(x)$, etc. The first term in (7) represents the answer to the question: who *benefits most* from an increase in community quality? If $F_{hL} > 0$, it is the better educated family, hence an efficiency gain from stratification. The second term measures whether an increase in L is more valuable to the average resident, when starting from a low or from a high level; if $F_{LL} < 0$, the *marginal productivity* of community quality is decreasing, implying an efficiency loss from stratification. The last term in (7) represents the

¹²There, we also explain how economy-wide complementarities can translate aggregate results into Pareto comparisons, and how welfare might differ between the short and long run.

¹³Naturally, if S changes curvature, the optimum may involve partial segregation.

answer to the question: where does a marginal well-educated agent, or family, *contribute most* to raising the quality of its community? If $L'' < 0$, such a family is much less valuable in the rich community which it joins than in the poor one which it leaves behind, leading to another loss from stratification.

In the basic model, agents allocate themselves solely on the basis of the first term in (7), no matter how small the private gains which it represents.¹⁴ More generally, wealth constraints and capital market imperfections also shape the residential equilibrium.

Proposition 6 *Agents segregate or integrate depending on $F_{hL} \gtrless 0$, no matter how small, and on other factors unrelated to the productivity of education. The equilibrium can be very inefficient, if the “concentration effects” F_{LL} or L'' are large in absolute value and have the opposite sign of F_{hL} , or more generally of the incentive to segregate R_{hx} .*

We shall generally be more interested in the case where the city stratifies, and the resulting concentration of human wealth on one side and poverty on the other is potentially inefficient – as Table 1 would tend to suggest. But Proposition 6 also makes clear when *inefficient mixing* will occur.

4 Stratification and inequality

Does families’ tendency to segregate into homogeneous communities make inequality more persistent? The simple answer is positive, as stratification compounds disparities in educational inputs at the family and community levels. The result is greater income inequality than would have occurred if families had remained integrated. Assuming perfect segregation ($n = 1/2$), we have:

$$h'_A / h'_B = F(h_A, h_A) / F(h_B, h_B) > F(h_A, \bar{L}) / F(h_B, \bar{L}), \quad (8)$$

in the simple case where purchased inputs play no role. When they do, the inequality is reinforced if $E'(x) > 0$, which is the empirically relevant case. In Section 4.2 we make the model truly dynamic and show how segregation slows down convergence among different groups, and can even generate poverty

¹⁴Half of the gains $2(F_L(h_A, L) - F_L(h_B, L))L'$ represents the private gains from trade accruing to a pair of A and B agents who exchange places. The other half is external: a marginal rise in L is more valuable, *ceteris paribus*, in the community with more L -sensitive agents; see the expression for S' . Interestingly, the results of Brooks-Gunn et al. (1993) suggest that $F_{hL} > 0$ but $F_{LL}, L'' < 0$.

traps. But first we point out that this story is subject to an important qualification, which tends to be overlooked. More specifically, we discuss what are usually *implicit assumptions* about the appropriate measure of inequality and the cost at which the rich are able to exclude the poor.

4.1 A caveat: inequality in what, and at what price?

The usual approach, in both the theoretical and empirical literatures, is to track across families the earnings, education levels or professional status of successive generations, as we did in (8). But it is not obvious that this is the appropriate measure of inequality. Better community quality and better schools come at a cost, namely a higher rent $\rho^1 > \rho^2$ and/or tax bill (if $t_x > 0$). Taking into account not only the payoff but also the costs of educational investments leads one to compare children's *total wealth* (human capital plus bequest), or more generally the *utility* levels of parents. This may lead to a very different answer.

Proposition 7 Consider the general model of Section 1, and assume $V_{hx} \geq 0$, $V_{h\rho} \geq 0$, ensuring (9).¹⁵

1. Equilibrium stratification need not increase inequality between rich and poor families' utility levels. If $n < 1/2$, then $U^1(h_A) - U^2(h_B) > \bar{U}(h_A) - \bar{U}(h_B)$, where the left and right hand sides correspond respectively to the stratified and integrated equilibrium. But if $n > 1/2$, the inequality is reversed, and if $n = 1/2$, the ranking is ambiguous.

2. The same is true of inequality in children's total wealth $c' + h'$, when utility is linear.

This simple result is proved in the appendix. It raises the old question of whether *child inequality* or *family inequality* is the relevant concept (Stiglitz (1973)), but also some new and more practical ones.¹⁶ What is the *cost to the rich of seceding from the poor*? Who owns the assets, property rights or legal rights which allow stratification to occur and earn the rents which it may create? In our model, answering this question means identifying the recipients of the capital gains and losses which stratification generates on land in communities 1 and 2 respectively. Proposition 7 is derived under the "neutral" assumption that they are distributed evenly: all housing initially belongs to absentee landowners, or alternatively to the

¹⁵We use this stronger condition for simplicity. The same result can be shown in a model with a continuum of types and communities (as in Wheaton (1993)) using only (3).

¹⁶Interestingly, a similar indeterminacy can be demonstrated in a model where agents self-select through a trade-off between school quality and tax rates, rather than land rents (e.g., Fernandez-Rogerson (1992)). Even though there is no "leakage" of community resources to outsiders, stratification need not increase inequality in utilities.

city's residents but with no correlation between an agent's type and the location of her property ($\nu_i = \eta_i = 1/2, i = A, B$). A more realistic assumption might be that A agents own all the land: $\nu_A = \eta_A = 1/2n, \nu_B = \eta_B = 0$. Yet even then, stratification need not increase inequality in total wealth: capital gains in C^1 may or may not offset capital losses in C^2 . What is required is that the rich capture a larger share of the rents which their exodus creates, and/or that the poor be left holding a larger share of the corresponding losses: $\nu_A > \nu_B, \eta_B > \eta_A$.

What this points to is the role played in the generation and persistence of inequality by various collective practices and institutions which allow the rich to *capture the benefits* of their secession, by raising -even temporarily- the relative cost to the poor of joining or remaining in a community which is appreciating (see Section 2.4). Prime among these is racial discrimination, whether enforced by law or custom. As long as Blacks are simply not allowed to bid for land in the suburbs to which White families move, the latter can regroup without dissipating too much of the resulting rents on higher land values. When, later on, Blacks are allowed in, those who do come must pay the full value of the community's social capital. Although the education gap between their children and those of their white neighbors will close, the gap in total wealth will not. In fact, the inequality in wealth which occurs the moment *de jure* segregation is lifted and property values adjust can in itself be sufficient to sustain *de facto* economic segregation. Consider the specification of Section 2.3, with $n = 1/2$. Suppose White and Blacks have the same amount z of total non-land wealth, and each own their houses in C^1 and C^2 respectively. Proposition 4 implies that the two groups will remain separated even once free mobility is allowed, though a rent differential $\rho^1 - \rho^2$ such that $\tilde{u}(z + \rho^1 - \rho^2) - \tilde{u}(z) < w(F(h_A)) - w(F(h_B)) < \tilde{u}(z) - \tilde{u}(z + \rho^2 - \rho^1)$. In recent years, the main device used to "keep out the poor" has been the right given to a community's residents to impose zoning regulations, which essentially operate as minimum income requirements (Durlauf (1992), Wheaton (1993), Fernandez and Rogerson (1993a)).¹⁷

The *empirical implication* of these observations is that studies which attempt to measure the role of peer effects, neighborhood spillovers or school quality in the intergenerational persistence of inequality, should try to take into account not only labor earnings but also financial wealth: how much of the human

¹⁷Equivalently, Durlauf (1993) assumes that a neighborhood's land price adjusts to the (equilibrium) value which keeps out the marginally poorer family, only once all richer families have moved in and bought the land at cost. In all these models, it still remains that some of the higher human capital passed on to children in rich communities reflects the higher taxes paid by their parents.

wealth afforded to a child by a better educational environment is reflected in a reduction of the bequest she receives? Answering this question might also help determine whether the peer effects found by Borjas (1992a) among members of the same ethnic group truly represent “ethnic capital”, or whether they really operate through local spillovers in ethnically segregated neighborhoods –as Borjas (1992b) would seem to indicate. While the former are inherited “for free” through birth, the latter are subject to choice and carry a price in the form of rent or tax differentials.

Bearing in mind this section’s caveat about the links between stratification and different measures of inequality, we now extend the analysis to a dynamic setting. We demonstrate how stratification gives rise to ghettos or poverty traps, and how those can in turn adversely affect overall growth.

4.2 Overlapping generations

Let the two periods considered until now represent any two overlapping generations. To keep things simple, the distribution of human wealth must be the only relevant state variable, determining both residential equilibrium and the appropriate measure of inequality. The first condition holds if utility is linear and capital markets are perfect (see Section 2.1 more generally), or alternatively if there are no financial bequests: $\omega(h)$, $y(h)$ represent labor income in the two periods of life, and all of c' is consumed in old age. The second condition requires that either inequality in children’s earnings be the criterion of interest, or that the first generation of A ’s who secede capture the whole (present) value of rents created in community 1, leaving the B ’s holding the less valuable land in community 2. Thus all housing is owner-occupied, and human capital inequality translates directly into inequality of total wealth or utility.

For simplicity we abstract from purchased local public goods, so that $h' = F(h, L)$. Finally, let $n = 1/2$; thus whether integrated or segregated, the two groups always remain homogeneous. Also, in either case, if $h_{A,t} > h_{B,t}$ the same is true for the next generation. Therefore if (3) is satisfied for all (h, ρ, x) , all generations of rich and poor segregate. The resulting impact on the persistence of inequality is most clear when $F(h, L) = \theta h^\alpha L^\beta$, with $\alpha < \alpha + \beta \leq 1 < \theta$. In the integrated equilibrium, the level of inequality $\chi_{t+1} \equiv \log(h_{A,t+1}/h_{B,t+1}) = \alpha \chi_t$ converges to zero at the rate $1 - \alpha$. In the stratified equilibrium, $\chi_{t+1} \equiv \log(h_{A,t+1}/h_{B,t+1}) = (\alpha + \beta) \chi_t$; initial conditions vanish at a slower rate, or even not at all if

$\alpha + \beta = 1$.¹⁸

Inequality can be even more persistent if $F(h, L)$ has decreasing returns in h alone, but increasing returns in (h, L) jointly over some range. Such local increasing returns, *realized only through segregation*, are what underlies ghettos or local poverty traps (Bénabou (1993), Durlauf (1992), (1993), Lundberg and Startz (1993)). Suppose for instance that local spillovers involve a threshold effect:

$$F(h, L) = \theta h^\alpha \max\{L - f, 0\}^{1-\alpha} \quad (9)$$

where $f > 0$ and $L(x) = (h_A)^x (h_B)^{1-x}$ is defined as the geometric average. Here all human capital magnitudes h, L, F , etc., are measured as deviations from some fixed, basic level $\underline{h} > 0$; thus $h' = 0$ is not to be taken literally. When the two classes are segregated, we have $h_{i,t+1} = F(h_{i,t}, h_{it})$, $i = A, B$. When they remain mixed, we have $h_{i,t+1} = F(h_{i,t}, L_t)$; given (9), this implies $L_{t+1} = F(L_t, L_t)$ since L_t is the geometric average. Hence the following result, which is illustrated in Figure 2.

Proposition 8 *Let human capital accumulation be given by $h_{i,t+1} = F(h_{i,t}, L_t^j)$, as defined in (9), and let $f < h_{B,0} < h^* \equiv f/(1 - \theta^{-1/(1-\alpha)}) < \bar{L}_0$. Under integration, inequality converges to zero, and both $h_{B,t}$ and $h_{A,t}$ grow asymptotically at the rate $\theta - 1 > 0$. Under segregation, the human capital poor families converges to \underline{h} in finite time, while that of rich families grows asymptotically at the rate $\theta - 1$.*

In equilibrium, stratification occurs (due to $F_{hL} > 0$ or the other factors discussed earlier) confining poorer families to a steady state of low productivity and income, which cohabitation with better educated dynasties would have allowed them to escape.¹⁹ Two recent papers provide supporting evidence for the kind of non-convexity which underlies the ghetto phenomenon. Crane (1991) finds threshold effects in the way a neighborhood's professional mix influences local rates of high-school dropout and teenage pregnancy. Using a methodology specifically suited to detect non-linearities in income dynamics, Cooper, Durlauf and Johnson (1993) find that very low levels of mean county income generate poverty traps.

¹⁸Estimating such a log-linear relationship for the descendants of immigrants to the United States, Borjas (1992a) (1992b) finds that human capital spillovers slow down convergence markedly: α and β both equal about .25 and are statistically significant.

¹⁹Tamura (1991b) and Galor and Tsiddon (1992) present models which combine increasing returns at the individual (country or family) level with an economy-wide technological externality, through which the poor eventually catch up with the rich. The implicit assumption is that the rich do not segregate (i.e., limit interactions to their own group), as they do when given the opportunity in our model.

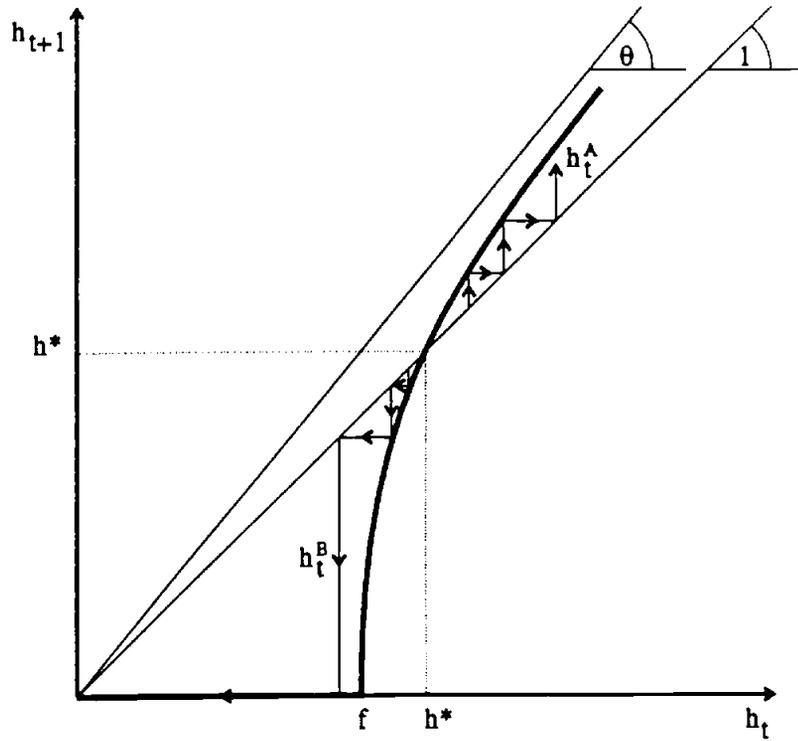


Figure 3a Stratification: $h_{t+1}^i = F(h_t^i, h_t^i)$

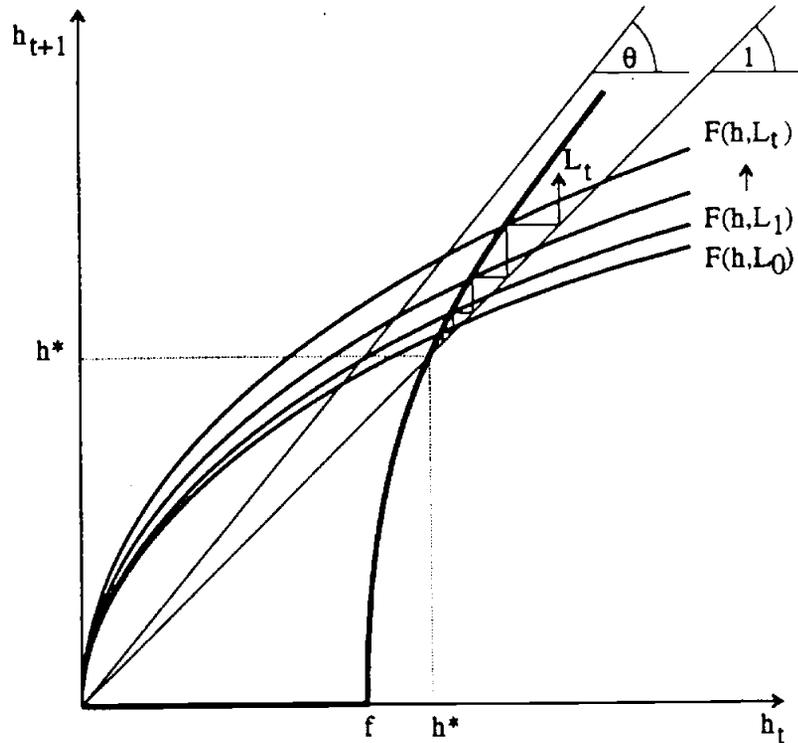


Figure 3b Integration: $h_{t+1}^i = F(h_t^i, L_t)$, $L_t = (h_t^A)^{1/2} (h_t^B)^{1/2}$

Stratification may even hurt the productivity and income of richer dynasties, when, in addition to *local* interactions in education, the different classes interact at the *global*, or economy-wide level. The simplest way in which city and suburban residents can be bound together is as complementary factors in the production of goods or knowledge. Let for instance the more educated *A* agents be managers, and the less educated *B* agents workers. A representative firm's output is $Y = \left(H_M^{\frac{\sigma-1}{\sigma}} + H_W^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$, where $H_M = n h_A$ is managerial human capital and $H_W = (1 - n) h_B$ worker human capital. The resulting incomes, $y_A = h_A^{\frac{\sigma-1}{\sigma}} (Y / n)^{\frac{1}{\sigma}}$, $y_B = h_B^{\frac{\sigma-1}{\sigma}} (Y / (1 - n))^{\frac{1}{\sigma}}$ are clearly interdependent. Task assignments are exogenous, but a model with occupational choice and competitive labor markets would have very similar implications (Bénabou (1993)).²⁰ Instead of production complementarities, agents could be linked through economy-wide knowledge spillovers (Lucas (1988), Romer (1986), Tamura (1991b)) or nationally funded public goods. These ties can all be captured by writing individual productivity as:

$$y(h) = h^\lambda H^{1-\lambda} \tag{10}$$

where H is some economy-wide index of human capital. All variables are still measured as deviations from \underline{h} , which represents a basic level of skills and earnings. Like that of the local index L , the sensitivity of H to heterogeneity is a key determinant of the costs and benefits of stratification. When H is a CES index, this sensitivity reflects the degree of complementarity among agents' skill levels; see Bénabou (1992). Here we shall simply assume that H , like L , is a geometric average: $H = (h_A)^n (h_B)^{1-n}$.

Proposition 9 *Assume the same conditions as in Proposition 8, and the production technology (10). Under integration, inequality converges to zero, and all earnings grow asymptotically at the rate $\theta - 1$. Under segregation, human capital inequality keeps rising, but all earnings converge to \underline{h} in finite time.*

In this economy, productivity growth slows down and eventually peters out because the secession of "managerial" dynasties prevents "worker" dynasties from acquiring the skills necessary to keep up.²¹

²⁰Alternatively, Tamura (1991a) and Bénabou (1992) present models where increasing returns cause agents to specialize in imperfectly substitutable intermediate inputs. The induced mapping from human capital to income is essentially the same as the one below.

²¹The results of Glaeser, Scheinkman and Shleifer (1993) confirm the importance of a well educated general labor force, in contrast to a high concentration of human capital on a small elite. In explaining a city's growth, they find the proportion of residents with 12 to 15 years of education (high school graduates and some college) to be both more important and statistically more significant than the proportion with higher education.

Making those skills essential to production (by defining H as the geometric average, or Y as a Cobb-Douglas) is only a convenient shortcut. When education uses real resources (schools, R&D), as in the general model $h' = F(h, L, E)$, reductions in productivity feed back into each family's human capital accumulation. Thus even when the elasticity of substitution σ between h_A and h_B in H is greater than one, and even in the absence of threshold effects, stratification can lower the steady-state path of human capital and income to which rich families converge. Whether or not the secession of the rich is ultimately *self-defeating* depends in particular on the relative costs of local and aggregate heterogeneity, measured respectively by the degrees of complementarity $1/\epsilon$ and $1/\sigma$ in local and global interactions.

As shown by Proposition 9, society may be faced with an *intertemporal tradeoff*. In the short run, stratification benefits the rich more than it hurts the poor, so that overall growth increases ($S'' > 0$ in Section 3). But over the long run, excessive heterogeneity acts as a drag on growth, affecting all lineages. There is then an incentive for the rich, provided their intergenerational discount rate is low enough, to subsidize the education of the poor. They could for instance agree to inter-community transfers (Cooper (1992)) or support switching to a system of national funding (Bénabou (1992)). But as shown by the fact that purchased inputs play no role in Proposition 9, these may be poor substitutes for the actual integration of schools, neighborhoods, networks, and so on. We return to this issue when discussing policy later on.

5 School financing

We now explicitly incorporate local school expenditures and taxes into the analysis. We show how they represent an additional segregating force, then examine the consequences for efficiency and inequality. These results will also form the basis for the policy evaluation conducted in the next section.

5.1 The choice of education expenditures

We start with the most simple case, where agents are risk-neutral, capital markets frictionless, and local schools financed by lump-sum or property taxes (with inelastic lot size, these are equivalent). This implies that each family's tax bill equals per-student expenditures E , and that a community's educational investment is not constrained by its tax base. Section 5.4 considers income tax financing, and Section 5.5 shows how a simple adaptation of the model can incorporate wealth constraints. We assume that in a community

with composition $(x, 1 - x)$ and spillover quality L , education expenditures are determined as:

$$E^*(x, L) \equiv \arg \max_E \{ xF(h_A, L, E) + (1 - x)F(h_B, L, E) - (1 + r)E \}. \quad (11)$$

We choose this decision rule for two reasons. First, unlike majority voting, it reflects in a simple, continuous manner both the proportions and the preferences of local residents. It can in fact be interpreted as the outcome of unrestricted vote-trading.²² Second, it yields the efficient outcome, *conditional* on community composition; this allows us to highlight the issue of *inter-* rather than *intra-*community efficiency. Equation (11) implies:

$$\begin{aligned} \frac{\partial E^*(x, L)}{\partial x} &= \frac{F_E^A - F_E^B}{-xF_{EE}^A - (1-x)F_{EE}^B} \approx \frac{F_{hE}}{-F_{EE}} \cdot \Delta h, \\ \frac{\partial E^*(x, L)}{\partial L} &= \frac{x F_{EL}^A + (1-x) F_{EL}^B}{-xF_{EE}^A - (1-x)F_{EE}^B} \approx \frac{F_{EL}}{-F_{EE}}, \end{aligned}$$

where F_{hE}^i , $i = A, B$ stands for $F_{hE}(h_i, L, E^*(L, x))$, etc. The approximations hold when $\Delta h/h$ is small, with all derivatives evaluated at $(\bar{h}, L(\bar{n}), E(\bar{n}))$. Although this condition is not required for any of the results, it simplifies notation and makes intuitions more transparent. The school budget $E(x) = E^*(x, L(x))$ varies with community composition as:

$$E'(x) = \frac{F_E^A - F_E^B + (x F_{EL}^A + (1-x) F_{EL}^B) L'(x)}{-xF_{EE}^A - (1-x)F_{EE}^B} \approx \frac{F_{hE}}{-F_{EE}} \cdot \Delta h + \frac{F_{EL}}{-F_{EE}} \cdot L' \quad (12)$$

When E and L are complements, rich communities may thus spend more on their students than poor ones, even though, were they placed in the same environment, students from poor backgrounds would have a higher marginal product of education expenditures than those from richer families ($F_{hE} < 0$).

5.2 Stratification

Having determined $E(x)$ and $t(h, \rho, x) = E(x)$, we now replace them in the bid-rent differential $R_{hx}(h, \rho, x)$, or equivalently in $SEG \equiv (1 + r)(R_x(h_A, \rho, x) - R_x(h_B, \rho, x)) = (F_L^A - F_L^B) L' + (F_E^A - F_E^B) E'$. This expression defines the net private gains, in terms of children's human capital, to a marginal pair of A and B families who switch places during the stratification process:

²²A political process resulting in a weighted average of each group's preferred level, $E^*(x, L) = \beta(x) E^A(L) + (1 - \beta(x)) E^B(L)$, $\beta'(x) > 0$, would have implications very similar to those of (11). Note that we assume $F_{EE} < 0$ throughout. Also, the case of most interest is when the equilibrium involves a high degree of stratification; as stressed by Tiebout (1956), there is then near-unanimity in each community.

$$\begin{aligned}
SEG &= \left(F_L^A - F_L^B + (F_E^A - F_E^B) \frac{x F_{EL}^A + (1-x) F_{EL}^B}{-x F_{EE}^A - (1-x) F_{EE}^B} \right) L' + \frac{(F_E^A - F_E^B)^2}{-x F_{EE}^A - (1-x) F_{EE}^B} \\
&\approx \left((F_{hL} + \frac{F_{hE} F_{EL}}{-F_{EE}}) L' + \frac{(F_{hE})^2}{-F_{EE}} \Delta h \right) \Delta h.
\end{aligned} \tag{13}$$

The term in brackets represents the *complementarity*, both direct and indirect through E , between parental education and community quality. The last term, always positive, captures the greater incentive of each type of agent to join a community where individuals with *similar preferences*, being more numerous, have greater weight in the setting of policy. The first implication of (13) is that stratification always occurs in the absence of local externalities ($F_L = 0$). The second one is that local funding of education may *induce* stratification which would otherwise not have occurred ($F_{hL} < 0 < SEG$).²³ This is probably an important piece of the explanation for the difference in the extent of stratification between the United States and most other industrialized countries.

5.3 Efficiency

Recall that $E(x)$ is chosen optimally, given x . The efficient (surplus-maximizing) residential allocation is thus again determined by the concavity of net community output, now defined as:

$$S(x) \equiv x F(h_A, L(x), E(x)) + (1-x) F(h_B, L(x), E(x)) - (1+\tau) E(x) \tag{14}$$

From (11) and the envelope theorem, $S' = F^A - F^B + x F_L^A + (1-x) F_L^B$. Then, using (12) yields:

$$\begin{aligned}
S'' &= 2(F_L^A - F_L^B) L' + (x F_{LL}^A + (1-x) F_{LL}^B) (L')^2 + (x F_L^A + (1-x) F_L^B) L'' \\
&\quad - (x F_{EE}^A + (1-x) F_{EE}^B) (E')^2
\end{aligned} \tag{15}$$

The first three terms are due to local spillovers, which affect education both directly and through their interaction with E . Their interpretation is the same as in (7). The last term, always positive, arises from the optimal adjustment of expenditures to community composition x and quality L . Absent local spillovers (and imperfections, say, in capital markets), the tendency to segregate in response to differing preferences over education promotes efficiency (Tiebout (1956)). In general, however, it may have the opposite effect:

²³De Bartolome (1990) considers the case where $F_{hL} < 0 = F_{EL} < F_{hE}$ and where these two opposing effects result in an asymmetric but incompletely stratified equilibrium. Although his political mechanism is different, this can be understood as interior solution $n < x^1 < \min\{2n, 1\}$ to the equation $\int_{2n-x^1}^{x^1} SEG(x) dx = F(h_A, L(x^1), E(x^1)) - F(h_A, L(x^2), E(x^2)) - [F(h_B, L(x^1), E(x^1)) - F(h_B, L(x^2), E(x^2))] = 0$.

Proposition 10 *The decentralization of school funding can trigger stratification which is inefficient: it suffices that $F_{hL} < 0 < SEG$ and $S'' < 0$.*

Indeed, we can rewrite:

$$S'' = SEG + (F_L^A - F_L^B) L' + (x F_{LL}^A + (1-x) F_{LL}^B) (L')^2 + (x F_{LE}^A + (1-x) F_{LE}^B) L' E' + (x F_{LE}^A + (1-x) F_{LE}^B) L' E' \quad (16)$$

The term SEG represents the net *private gains* from stratification; the elements which determine its sign were discussed in (13). All others terms constitute *wedges* between social and private values. The first three of those involve only the local spillover, and were discussed in Section 3. The last one involves the interaction of E and L . As agents congregate into more homogeneous groups, they ignore they own effect on community quality L , including the resulting impact on the productivity of E . Assuming, realistically, that $E'(x) > 0$, the effect is favorable if L and E are complements: a planner would also like, *ceteris paribus*, for L to be high in the community where E is high. It is detrimental if they are substitutes.²⁴

5.4 Income taxes

We now consider income taxes: $t(h, x) = \tau(x) h$. We continue to assume that in each community, school expenditures $E(x)$ are chosen efficiently with respect to local surplus, that is, according to (11). The income tax rate is then simply determined by the balanced budget constraint: $\tau(x) = E(x) / A(x)$, where $A(x) = x h_A + (1-x) h_B$ is average community income. This leads to the following result, proved in the appendix.

Proposition 11 *Let the local spillover L be any CES index, and assume that $\Delta h / \bar{h}$ is small. If the education production function $F(h, L, E)$ is homogeneous of degree no greater than one, or more generally if the return to marginal expenditures decreases with the scale of educational inputs ($F_E(\lambda h, \lambda L, \lambda E) < F_E(h, L, E)$ for $\lambda > 1$), the school budget rises less than one for one with average community income. Therefore the income tax rate $\tau(x)$ decreases with x , promoting stratification.*

²⁴Equation (16) readily yields De Bartolome's (1990) result that an incompletely stratified equilibrium is inefficient, given his assumptions $F_{hL} < 0$, $F_{LL} < 0$ and $L'' = F_{EL} = 0$. As seen in the previous footnote, such an equilibrium is defined by $\int_{2n-x^1}^{x^1} SEG(x) dx = 0$; this implies $S'(x^1) - S'(2n-x^1) = \int_{2n-x^1}^{x^1} S''(x) dx < 0$, violating optimality.

5.5 Wealth effects and borrowing constraints

When agents have concave utility or face imperfect capital markets, their preferences over education spending reflect not only its contribution to their child's human capital, but also their different opportunity costs of funds. The effect of equilibrium land rents on wealth makes this case more complicated, but it can still be dealt with, as in De Bartolome (1990). There is, however, a simpler way to incorporate its essential implication, namely that communities with poorer residents have a lower ability to invest, due to their lower tax base. Following Becker (1975), let rich and poor agents have different but constant opportunity costs of funds, $r(h_A) < r(h_B)$.²⁵ Then, simply replace $F(h, L, E)$ by $\tilde{F}(h, L, E) \equiv F(h, L, E)(1 + r)/(1 + r(h))$ in the derivations of $E'(x)$ and $SEG(x)$; $r'(h) < 0$ then implies that poorer communities invest less than rich ones, other things equal. It also tends to make $\tilde{F}_{hL} > 0$, promoting stratification, as shown earlier. In the efficiency analysis the *actual* products $F(h, L, E)$ should be used (hence another wedge between $SEG(x)$ and $S''(x)$), unless the planner cannot provide loans more efficiently than the market, or seeks to maximize the sum of agents' utilities rather than net output. In such cases she should internalize agents' different costs of funds and the resulting gains from stratification, and use $\tilde{F}(h, L, E)$ to compute social surplus.

6 Policy implications

6.1 Background

In the United States, an increasing number of states are being forced by court decisions to close the wide gaps which exist between the school budgets of different communities. The underlying view is that these disparities, stemming from the lower property tax base of poorer towns, do not give equal starting chances to their young. Equalizing transfers, or state control of education funding, are then needed to reduce inequality as well as possible inefficiencies.²⁶

There is also a contrarian view, of which a recent article on Kansas City in *The Economist* (1993) is representative. Under a 1986 court order to remedy racial segregation in its schools, the city "scrapped its existing school system... and replaced it with the best... money could buy," at the cost of an extra \$1.3

²⁵More specifically, let $r(h)$ be a decreasing function of initial resources $\omega(h)$, so that wealth redistributions do affect educational choices.

²⁶The most recent example is Michigan, where voters recently approved a proposal to replace local property taxes by an increase in the state sales tax, as the primary source of funding for schools.

billion, or \$36,111 per student over the normal school budget. Three quarters of the cost were born by the state, resulting in severe cutbacks in other districts.

“So far, however, all this lavish expenditure has produced few of the desired results . . . The racial balance in the schools is the same today as it was when the order was issued. The past six years have seen no improvement whatsoever in children's scores in standardized tests of reading and math . . . Pupils in elementary schools which have not been turned into magnet schools regularly outperform pupils in generously funded magnet schools . . . The drop-out rate has risen every year, without fail, since the decree was handed down, and now stands at a disgraceful 60%.”

Not surprisingly, Kansas City school officials vigorously dispute this bleak assessment. But other pieces of evidence appear consistent with the view that lack of financial resources is not the main impediment faced by mediocre or failing schools. Expenditures per student are higher in some urban school districts than in neighboring suburban or rural ones, thanks to a larger industrial tax base or to state subsidies. The latter nonetheless achieve better test scores and dropout rates. Reviewing the empirical literature on education production functions, Hanushek (1986) observes that most studies fail to demonstrate any link between school expenditures and student achievement. Perhaps most importantly, the increased centralization of education funding experienced by many states (starting with the 1971 California Supreme Court “*Serrano*” decision) appears to have significantly reduced *spending* inequalities, but not *achievement* inequalities, as measured from SAT scores (Downes (1992)). Moreover, it seems to have been accompanied by a concomitant *decline* in those states' *average* test scores (Pelzman (1992)).

Opponents of the “lack of resources” hypothesis often point instead to culprits such as teachers' unions, the education bureaucracy, or the disincentive effects of the welfare system. We shall advance a very different explanation of the evidence, based in particular on complementarities between purchased and social inputs in education. We identify in particular a case where equalizing school expenditures fails to reduce segregation and leads to a decline in overall academic performance, hurting the rich more than it helps the poor. Yet a better designed policy may still improve *both* equity and efficiency.

6.2 Equalizing school budgets

We model equilibrium prior to state intervention using the simple specification of education expenditures introduced in Section 5.1: $E^j = E^*(x^j, L(x^j))$, where x^j is the fraction of high human capital families in community j . We thus deliberately abstract from financing constraints at the individual or community level. This is not because we consider the credit market imperfections and wealth effects discussed earlier as unimportant; they played a central role in our previous analysis of local and centralized school funding (Bénabou (1992)), as well as that of Fernandez and Rogerson (1993b).²⁷ What we seek to do here is to highlight the *limits* of redistributive policies, and explore whether there remains a role for the state even absent such constraints.

We first mention some ways in which attempts to equalize school budgets might be defeated at the local level. Clearly, outright redistributions of wealth leave E^1 and E^2 unaffected; only consumption adjusts. Mandating that E^1 be reduced and E^2 increased may induce creative accounting. A rich town might shift, say, athletic facilities or a library from the school budget to some other line item. A poor one might claim as education-related some expenditures which have mostly consumption value (nicer buildings, teacher and administrator salaries). Finally, taxing E^1 and rebating the proceeds to schools in community 2 as a block grant does reduce spending in C^1 , but changes nothing in C^2 , where residents simply reduce their contribution to education by the amount of the transfer. As a result, *total expenditures fall* (as observed by Downes (1992) for California following *Serrano*), and so do *test scores*.

But let us now take as given that the state is successful in enforcing equalization of school expenditures, either through matching grants (Feldstein (1975)) or simply by centralizing funding.²⁸ We start from a stable equilibrium where the population is stratified ($SEG > 0$ in (13)) and the richer town spends more on education ($E'(x) > 0$ in (12)). After state intervention, E^1 is reduced, and E^2 increased, to a common value \bar{E} .²⁹ There are two cases to consider.

²⁷Credit constraints were also at the center of earlier models where education was *privately* purchased, such as Loury (1977) or Glomm and Ravikumar (1992).

²⁸Between 1960 and 1983, the combined State and Federal share in elementary and secondary school spending rose from 35% to 50%, with a corresponding decline in reliance on local and private resources (Hanusheck (1986)).

²⁹ \bar{E} could be either the mid-point $(E^1 + E^2)/2$, or the optimal value $E(n) = E^*(n, L(n))$ which would be chosen by integrated communities, or the optimal value given the stratified configuration.

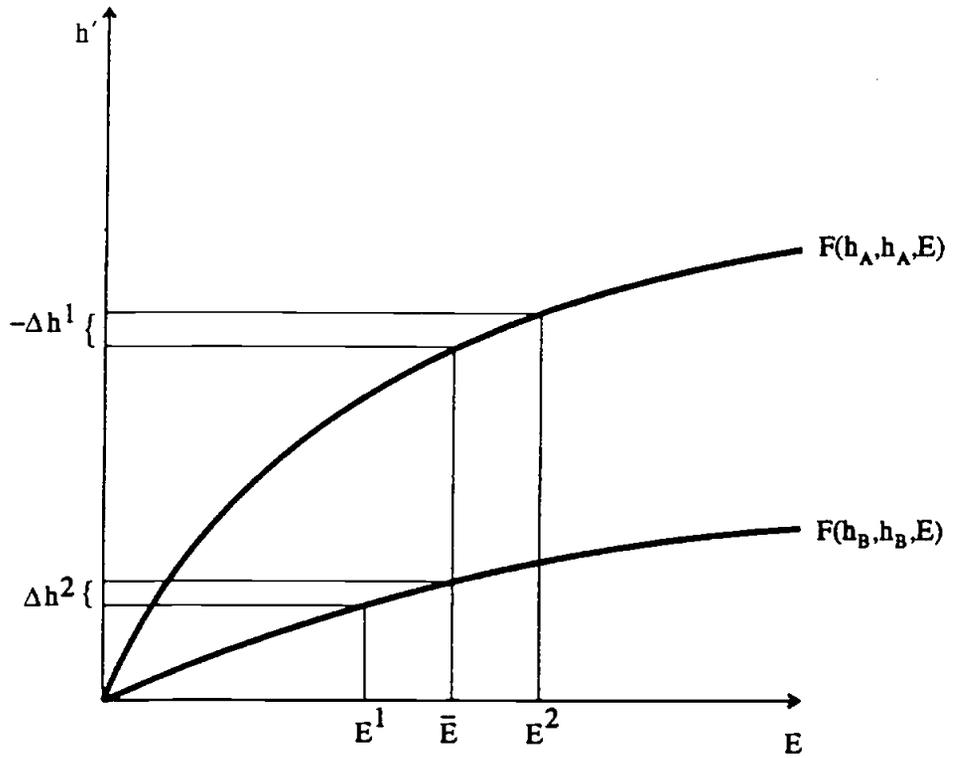


Figure 4a stratification persists

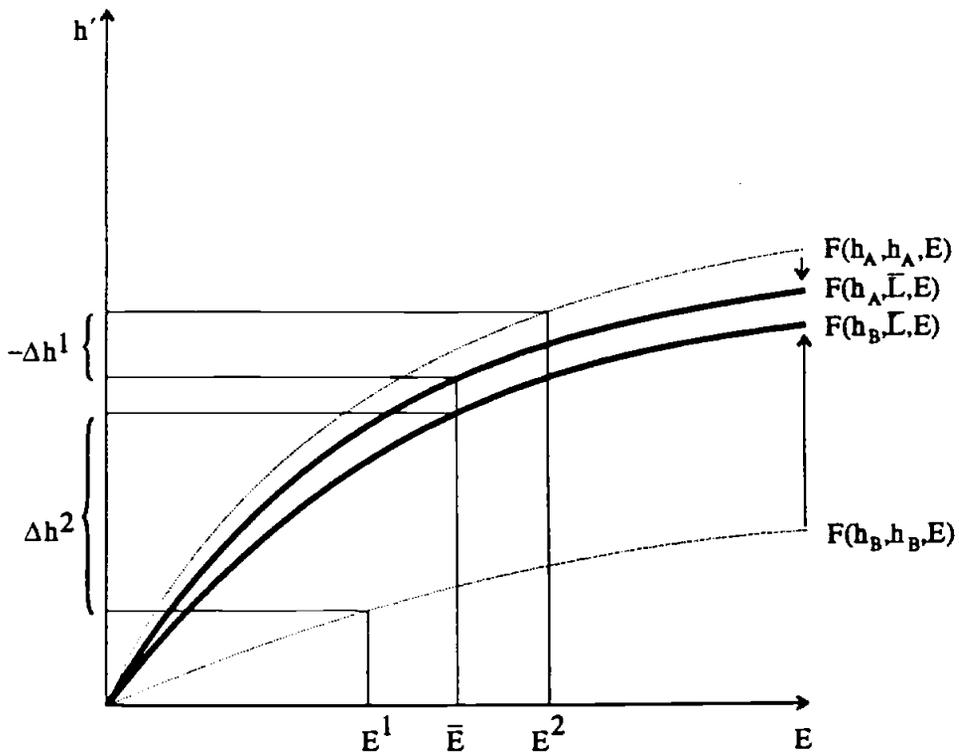


Figure 4b integration occurs

Case 1. Segregation persists

If $F_{hL} > 0$, the composition of each community remains unchanged. This could also be due to the other stratifying forces discussed earlier, particularly some group's preference for racial separation. Conditional on each x^j , expenditures were previously set optimally, so state intervention necessarily reduces surplus. What is interesting is the way in which this occurs, as illustrated by **Figure 3a** for the case of perfect stratification ($n = 1/2$). In C^1 , students' human capital is reduced by $-\Delta h^1 = x^1 (F(h_A, L^1, E^1) - F(h_A, L^1, \bar{E})) + (1 - x^1) (F(h_B, L^1, E^1) - F(h_B, L^1, \bar{E})) > (1 + r) (E^1 - \bar{E})$, due to the concavity of the maximization problem (11). In C^2 , the corresponding increase is $\Delta h^2 = x^2 (F(h_A, L^2, \bar{E}) - F(h_A, L^2, E^2)) + (1 - x^2) (F(h_B, L^2, \bar{E}) - F(h_B, L^2, E^2)) < (1 + r) (\bar{E} - E^2)$, by the same argument.

Proposition 12 *When $F_{hL} > 0$, equalizing school budgets leaves community composition unchanged but reduces total surplus. Student achievement and subsequent earnings decline more in rich communities than they improve in poor ones (per dollar of cutback or subsidy). If state-wide expenditures are kept constant, $\bar{E} = (E^1 + E^2)/2$, average student performance declines.*

So while income inequality is reduced, it is less by lifting up the poor than bringing down the rich. The intuition is that school expenditures are more productive in a human capital-rich community, due to complementarity either between h and E , or more interestingly, between E and L ; see (12).³⁰ Top-notch teachers, computers and other educational resources may not do much in a school plagued by discipline problems, the lack of motivation from role models in the community, peer pressure not to study, gangs, etc. Redistributing expenditures without simultaneously “redistributing” social capital is ineffective.

Case 2. Segregation is undone

If $F_{hL} \leq 0$ (and the other stratifying factors are not too strong), equalizing expenditures removes the incentive to segregate. The stable equilibrium then shifts from the stratified to the integrated configuration. Both purchased and non-purchased inputs E and L into education are equalized across students, achieving a greater reduction in inequality than in the previous case. There remains the question of whether academic

³⁰Downes and Pogue (1993) provide evidence that communities with a higher fraction of students from disadvantaged backgrounds have higher per-student costs, for a given level of educational achievement. While one must be careful about identifying average and marginal costs, this suggests that school expenditures in these communities have a lower marginal product.

achievement, income and the surplus from education rise more or less for poor students than they fall for rich ones. The answer is easily obtained in the natural case where \bar{E} is set at the optimal level for an integrated community, $\bar{E} = E^*(n, L(n))$. The net change in surplus is then $S(n) - (S(x^1) + S(2n - x^1))/2$, where S is given by equation (15); hence the following results, proved in the appendix.

Proposition 13 *When $F_{hL} < 0$, equalizing school budgets leads to integration. If state-wide expenditures are chosen optimally, $\bar{E} = E^*(n, L(n))$, total surplus increases provided S is concave. If, moreover, $\Delta h/\bar{h}$ is small and L is a CES index, surplus increases even when state-wide expenditures remain fixed, $\bar{E} = (E^1 + E^2)/2$. In that case average student achievement also rises.*

The conditions under which S is concave, so that integration improves both efficiency and equality, were analyzed earlier. This case is illustrated in **Figure 3b**: the equalization of L induced by the new policy shifts up the returns to educational expenditure on type B agents, much more than it shifts down the returns to expenditure on type A agents. The vertical difference between the average of the two solid curves and the two dashed ones is measured (to a second-order approximation), by the sum of the first three terms in (15). The distortions from constraining expenditures to be the same across communities and individuals are captured by the remaining term.

6.3 Reversing segregation

Consider now the case where equalizing expenditures is not sufficient to prevent segregation. This could be due to any of the stratifying forces identified in Section 1; for convenience we continue to focus on the complementarity $F_{hL} > 0$. Note that integration may still yield the surplus gains described in Proposition 13 and illustrated by **Figure 3b**; all that matters is that S be concave. In principle, a state government concerned with either inequality or inefficiency (if $S'' < 0$) could bring about integration by making poor communities sufficiently attractive to families with high human capital, and vice-versa. This could be done directly, through tax incentives or housing subsidies ($t_{hx}(h, x) > 0$) or indirectly, through a *contingent* allocation of educational resources $\hat{E}(x)$:

$$\widehat{SEG} = F_{hL}(h, L(x), \hat{E}(x)) L'(x) + F_{hE}(h, L(x), \hat{E}(x)) \hat{E}'(x) - (1+r)\hat{t}_{hx}(h, x) < 0, \quad (17)$$

for all x .³¹ The new stable equilibrium will be symmetric, with equal expenditures $\widehat{E}(n)$ everywhere; $\widehat{E}(n)$ can even be chosen optimally, to equal $E^*(n, L(n))$. But what this differential condition shows is that stratification is likely to be much *harder to undo* once it has occurred than it is to *stop in its tracks* early on. Due to the cumulative nature of the process, the amount of transfers required to induce the first few rich families to come back is considerably larger than what it would have taken to make them stay in the first place. For instance, when using only education expenditures, with $F_{AE} > 0$,

$$\widehat{E}(2n - x^1) - \widehat{E}(x^1) \gg \widehat{E}(n - \varepsilon) - \widehat{E}(n + \varepsilon),$$

where ε is small. Large transfers between the school budgets of rich and poor towns, even if only temporary, may well be politically infeasible. The dynamic aspects of such policies also raise problems of credibility. This kind of *irreversibility* may explain why magnet schools and similar programs recently implemented in some US cities are not very successful at fostering racial and economic desegregation of schools: the scope of redistribution required for such policies to be effective (advocates speak of a “Marshall Plan” for the inner cities) seems to be politically unacceptable. In metropolitan areas which are new or have retained a balanced urban-suburban composition, on the other hand, modest policy interventions can be effective; see Rusk (1993). But the point is perhaps most relevant for European countries where despite a centralized, fairly egalitarian allocation of school resources, economic and racial stratification appears to be on the rise, with quasi-ghettos starting to emerge at the periphery of many cities. If action is to be taken, it should be before polarization has reached the point where large gains and losses have become locked in.

7 Conclusion

The model developed in this paper allowed us to link together important issues of local public finance, income distribution and productivity growth. Five main results emerge. First, minor differences in education technologies, preferences, or wealth can lead to a high degree of stratification. Imperfect capital markets are not necessary, but will compound these other sources. Second, stratification makes inequality

³¹An important issue, not taken up here, is whether competition among communities allowed to charge different prices $t(h, \rho, x)$ to agents with different characteristics will also achieve efficiency; see Eden (1992) and Epple and Romano (1993) for an analysis of school competition with student vouchers. The use of E to improve on an inefficient residential equilibrium, which does not require family characteristics to be observable, was considered by De Bartolome (1990). It can only be implemented in a centralized manner.

in education and income more persistent across generations. Whether or not the same is true of inequality in total wealth depends on the ability of the rich to appropriate the rents created by their secession. Third, the polarization of urban areas resulting from individual residential decisions can be quite inefficient, both from the point of view of aggregate growth and in the Pareto sense, especially in the long run. Fourth, when state-wide equalization of school expenditures is insufficient to reduce stratification, it may improve educational achievement in poor communities much less than it lowers it in richer communities; thus average academic performance and income growth both fall. Yet policies better tailored to the interaction between purchased and non-purchased educational inputs may still improve both equity and efficiency. Fifth, because of the cumulative nature of the stratification process (hence of any offsetting transfers), it is likely to be much harder to reverse once it has run its course than to arrest at an early stage.

Important issues not addressed in the paper include the role of private schools and the effectiveness of voucher programs for education or housing. It would be interesting to try and extend the model to explore these matters as well. But perhaps the most intriguing direction for further research follows from two of the points mentioned above, concerning dynamics. The first is the creation and appropriation of rents during the "real time" process of stratification. The second is how considerations of credibility and political equilibrium constrain the time path of desegregation policy, or even its feasibility. These issues merit a more detailed analysis than the first steps taken here.

Appendix

A-Proof of Proposition 1. For all $(\rho^1, x^1, \rho^2, x^2)$ let $\Delta V(h \mid \rho^1, x^1, \rho^2, x^2) \equiv V(h, \rho^1, x^1) - V(h, \rho^2, x^2)$. Standard reasoning shows that if $R_{hx} > 0$ (respectively, < 0) everywhere, then $\Delta V(h) \geq 0$ implies $\Delta V(h') > 0$ for all $h' > h$ (respectively, $h' < h$), with strict inequality if $x^1 > x^2$. It is then straightforward to show that an equilibrium is a solution to one of the following three conditions.

(a) $\Delta V(h_A) = \Delta V(h_B) = 0$. This requires $x^1 = x^2 = n$, hence $\rho^1 = \rho^2$. Conversely, this symmetric allocation is always an equilibrium.

(b) $\Delta V(h_A) > \Delta V(h_B) = 0$. This requires $x^2 = 0$, $x^1 = 2n$, hence $2n \leq 1$. Clearly, one can always find ρ^1, ρ^2 such that $\Delta V(h_B \mid \rho^1, 2n, \rho^2, 0) = 0$. When $R_{hx} > 0$ (respectively, < 0), the other condition is automatically (respectively, never) satisfied, hence the allocation is (respectively, cannot be) an equilibrium.

(c) $\Delta V(h_A) = 0 > \Delta V(h_B)$. This requires $(1 - x^1)/2 = 0$, $(1 - x^2)/2 = 1 - n$, hence $2n \geq 1$. One can always find ρ^1, ρ^2 such that $\Delta V(h_A \mid \rho^1, 1, \rho^2, 2n - 1) = 0$. The rest of the proof is identical to case (b).

This concludes the derivation of equilibria. We now turn to stability, which is defined very simply. An equilibrium $(\rho^1, x^1, \rho^2, x^2)$ is said to be stable if, for all small $|\epsilon|$, moving ϵ A agents from C^2 to C^1 and the same number of B agents in the reverse direction makes B agents the highest bidders for land in C^1 , and vice versa. Formally:

$$\epsilon(\Delta V(h_A \mid \rho^1, x^1 + \epsilon, \rho^2, x^2 - \epsilon) - \Delta V(h_B \mid \rho^1, x^1 + \epsilon, \rho^2, x^2 - \epsilon)) < 0,$$

for any feasible $\epsilon \geq 0$. Clearly, the symmetric equilibrium is stable if and only if $R_{hx} < 0$. The stratified equilibrium, which exists only when $R_{hx} > 0$, is then stable since it satisfies $\Delta V(h_A) - \Delta V(h_B) > 0$, which is preserved by continuity for small perturbations $\epsilon < 0$ (which are the only feasible ones).

B-Alternative Wealth-Sorting Mechanisms. (1) Taxes. Consider first the case of lump-sum taxes or user fees. For instance, if $V(h, T, E) = \tilde{u}(h - T) + w(E)$, $-V_E / V_T$ clearly increases with h . As a standard result, agents segregate as much as permitted by the number of communities. Sorting through proportional taxes requires restrictions on preferences: the income effect must dominate the substitution effect. For instance, let $V(h, \tau, x) = \tilde{u}((1 - \tau)h) + w(\tau L(x))$, where $L(x) \equiv x h_A + (1 - x) h_B$ is average community income and $E(x) = \tau L(x)$ school quality, as in Fernandez-Rogerson (1992). Then $-V_x / V_\tau = \tau L'(x) w'(\tau L(x)) / (h \tilde{u}'((1 - \tau)h))$ increases with h if and only if \tilde{u} has relative risk aversion $-z \tilde{u}''(z) / \tilde{u}'(z) > 1$.

(2) *Status groups*. Let us re-label each community C^j a “professional occupation”, conferring on its members: (i) a social status which increases with the fraction x^j of highly educated (A) co-workers; (ii) a wage $\mu^j = \mu(H^j)$ per unit of human capital, which falls with the sector’s total human capital input H^j , due to diminishing returns. Agents derive utility from wages and from belonging to a higher status group: $V(h, \mu^j, x^j) = \tilde{u}(\mu^j h) + w(x^j)$. Under the same condition as above, $V_x / V_\mu = w'(x^j) / (h \cdot \tilde{u}'(\mu^j h))$ increases with h : better educated agents are more willing to accept a lower remuneration in exchange for higher status. As a result, the labor market stratifies, so that at least one occupation has a homogeneous workforce. The specification used here differs somewhat from Fershtman and Weiss (1992), but the basic mechanism and result are similar.

C-Proof of Proposition 7. (1) In the first case, some B agents remain in C^1 after stratification, so $U^1(h_A) - U^2(h_B) = U^1(h_A) - U^1(h_B) = V(h_A, \rho^1, x^1) - V(h_B, \rho^1, x^1) > V(h_A, \bar{\rho}, \bar{x}) - V(h_B, \bar{\rho}, \bar{x})$ since $x^1 > \bar{x}$ and $\rho^1 > \bar{\rho}$. The opposite reasoning applies when it is C^2 which remains mixed after agents sort themselves. Part (2) is immediate

D-Proof of Proposition 11. First, $\tau'(x) \leq 0$ if and only if $E'(x)A(x) \leq E(x)A'(x)$, or using (12):

$$\psi \equiv A(x)(F_E^A - F_E^B + (xF_{EL}^A + (1-x)F_{EL}^B)L'(x)) + (-xF_{EE}^A - (1-x)F_{EE}^B)\Delta h \leq 0$$

For small $\Delta h / \bar{h}$, $A(x) \approx \bar{h}$ and for any CES index, $\bar{L} \approx \bar{h}$, $L'(x) \approx \Delta h$, to a second-order approximation. Therefore $\psi \approx \Delta h(\bar{h}F_{hL} + \bar{L}F_{EL} + \bar{E}F_{EE})$, where $\bar{E} \equiv \arg \max_E \{F(\bar{h}, \bar{L}, E) - (1+r)E\}$ and all derivatives are evaluated at $(\bar{h}, \bar{L}, \bar{E})$. Finally, $g(\lambda) \equiv F_E(\lambda h, \lambda L, \lambda E) - F_E(h, L, E) < 0$ for all $\lambda > 1$ implies $g'(1) \leq 0$, hence the result. Note in particular that when $F(h, L, E)$ is homogeneous of degree one, $\tau'(x) = 0$.

E-Proof of Proposition 13. Only the case where $\bar{E} = (E^1 + E^2)/2$ remains to be proved. With expenditures fixed, the change in average student achievement equals the change in total surplus. This, in turn, equals the change in surplus achieved with the optimal $E^*(n, L(n))$, minus the distortion from choosing \bar{E} instead. By the envelope theorem, this distortion is of second order in $\delta \equiv \bar{E} - E^*(n, L(n)) = (E(x^1) + E(x^2))/2 - E((x^1 + x^2)/2)$. But δ is itself of second order in Δh , as can be seen from (12): when L is a CES with elasticity ϵ , $L' \approx \Delta h$ and $L'' \approx (\Delta h)^2/2\epsilon$; hence $E'(x)$ is of order Δh , and, differentiating (12), $E''(x)$ is of order $(\Delta h)^2$. The distortion δ^2 is therefore negligible compared to the gains $S(n) - (S(x^1) + S(x^2))/2$ which are of second order in Δh , as shown by the expression (15) giving S'' .

References

- Baumol, W. (1967) "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *American Economic Review*, 57, 415-426.
- Becker, G. (1975) "Human Capital: a Theoretical and Empirical Analysis with Special Reference to Education," Columbia University Press, New York.
- Bénabou, R. (1992) "Heterogeneity, Stratification, and Growth," MIT Working Paper 93-4, December.
- Bénabou, R. (1993) "Workings of a City: Location, Education and Production," *Quarterly Journal of Economics*, 108, 619-652.
- Borjas, G. (1992a) "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics*, 108, p. 123-150.
- Borjas, G. (1992b) "Ethnicity, Neighborhoods, and Human Capital Externalities," U.C. San Diego mimeo, September.
- Brooks-Gunn, J., Duncan, G., Kato Klebanov, P. Sealand, N. (1993) "Do Neighborhoods Influence Child and Adolescent Development?," *American Journal of Sociology*, 99, 353-395.
- Card, D. and Krueger, A. (1992) "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 100, 1-40.
- Case, A. and Katz, L. (1991) "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth," NBER Working paper No. 3705.
- Coleman, J. (1988) "Social Capital in the Formation of Human Capital," *American Journal of Sociology*, 94, S95-S120.
- Cooper, S. (1992) "A Positive Theory of Income Redistribution," Stanford University mimeo, July.
- Cooper, S., Durlauf, S. and Johnson, P. (1993) "On the Evolution of Economic Status Across Generations," University of Oregon Working Paper 146, October.
- Corcoran, M., Gordon, R., Laren, D., and Solon, G. (1990) "Effects of Family and Community Background on Community Status," *American Economic Review*, 80, 362-366.
- Cornes, R. (1993) "Dyke Maintenance and Other Stories: Some Neglected Types of Public Goods," *Quarterly Journal of Economics*, 108, 259-272.
- Crane, J. (1991) "The Epidemic Theory of Ghettos, and Neighborhood Effects on Dropping out and Teenage Childbearing," *American Journal of Sociology*, 96, 1226-1259.
- Datcher, L. (1982) "Effects of Community and Family Background on Achievement," *Review of Economic and Statistics*, 64, 32-41.
- De Bartolome, C. (1990) "Equilibrium and Inefficiency in A Community Model with Peer Group Effects," *Journal of Political Economy*, 98, 1, p. 110-133.
- Downes, T. (1992) "Evaluating the Impact of School Finance Reform on the Provision of Public Education: The California Case," *National Tax Journal*, 45 405-419.

- Downes, T., and Pogue, T. (1993) "Adjusting School Aid Formulas for the Higher Cost of Educating Disadvantaged Students," Northwestern University mimeo, June.
- Durlauf, S. (1992) "A Theory of Persistent Income Inequality," NBER Working Paper 4056, April.
- Durlauf, S. (1993) "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," University of Wisconsin-Madison Working paper mimeo, April.
- Dynarski, M., Schwab, R., and Zampelli, E. (1989) "Local Characteristics and Public Production: The Case of Education," *Journal of Urban Economics*, 26, 250-263.
- The Economist*, (1993) "The Cash Street Kids," August 28th-September 3^d, p. 23-25.
- Eden, B. (1992) "How to Subsidize Education and Achieve Voluntary Integration: An Analysis of Voucher Systems," University of Iowa mimeo, December.
- Epple, D. and Romano, R. (1993) "Competition Between Private and Public Schools, Vouchers and peer Group Effects," Carnegie-Mellon University mimeo, April.
- Feldstein, M. (1975) "Wealth Neutrality and Public Choice in Education," *American Economic Review*, 65, 75-89.
- Fernandez, R. and Rogerson, R. (1992) "Income Distribution, Communities and the Quality of Public Education: A Policy Analysis," mimeo, Boston University, July.
- Fernandez, R. and Rogerson, R. (1993a) "Keeping People Out: Income Distribution, Zoning and the Quality of Public Education," Boston University Working Paper No. 19, January.
- Fernandez, R. and Rogerson, R. (1993b) "Public Education and the Dynamics of Income Distribution: A Quantitative Evaluation of Education Finance Reform," Boston University mimeo, November.
- Fershtman, C., and Weiss, Y. (1992) "Social Status, Culture and Economic Performance," Tel-Aviv University mimeo, November.
- Galor, O. and Tsiddon, D. (1992) "Income Distribution and Output Growth: the Kuznets Hypothesis Revisited," Brown University mimeo, November.
- Galor, O. and Zeira, J. (1993) "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60, 35-52.
- Glaeser, E., Scheinkman, J. and Shleifer, A. (1993) "Economic Growth in a Cross-Section of Cities," Harvard University mimeo, April.
- Gloinn, G. and Ravikumar, B. (1992) "Public vs. Private Investment in Human Capital: Endogenous Growth and Income Inequality," *Journal of Political Economy*, 100, 818-834.
- Hamilton, B. and Macauley, M. (1991) "Determinants and Consequences of Private-Public School Choice," *Journal of Urban Economics*, 29, 282-294.
- Hanushek, E. (1986) "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature*, 24 1147-1177.
- Jencks, C. and Mayer, S. (1990) "The Social Consequences of Growing Up in a Poor Neighborhood," in *Inner City Poverty in the United States*, L. Lynn and M. McGreary, eds., National Academy Press,

Washington DC.

Kozol, J. (1991) *"Savage Inequalities: Children in America's Schools,"* Crown Publishers, New York.

Loury, G. (1977) "A Dynamic Theory of Racial Income Differences," in A. Le Mond (eds.) *"Women, Minorities and Employment Discrimination,"* Lexington Books, MA.

Loury, G. (1987) 'Why Should We Care about Group Inequality?', *Social Philosophy and Policy*, 5, 249-271.

Lucas, R. Jr. (1988) "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3-42.

Lundberg, S., and Startz, R. (1993) "On the Persistence of Racial Inequality," University of Washington mimeo, August.

Manski, C. (1993) "Identification of Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531-542.

Montgomery, J. (1991) "Social Networks and Persistent Inequality in the Labor Market: Toward an Economic Analysis," *American Economic Review*, 81, 1408-1418.

Miyao, T. (1978) "Dynamic Instability of a Mixed City in the Presence of Neighborhood Externalities," *American Economic Review*, 68, 454-463.

Pelzman, S. (1992) "The Political Economy of the Decline of American Public Education," Working Paper No. 78 (June), Center for the Study of the Economy and the State, University of Chicago.

Romer, P. (1986) "Increasing Returns and Long-Term Growth," *Journal of Political Economy*, 94, 1002-10037.

Rusk, D. (1993) *"Cities Without Suburbs,"* John Hopkins University Press, Baltimore

Sah, R. (1991) "Social Osmosis and Patterns of Crime," *Journal of Political Economy*, 99, 1272-1295.

Schelling, T. (1978) *"Micromotives and Macrobehavior,"* W.W. Norton and Company, New York.

Stiglitz, J. (1973) "Education and Inequality," *The Annals of the American Academy of Political and Social Sciences*, 409, 35-145.

Streufert, P. (1991) "The Effect of Underclass Social Isolation on Schooling Choice," University of Wisconsin-Madison Working Paper DP 954-91, August.

Tamura, R. (1991a) "Efficient Equilibrium Convergence, Heterogeneity and Growth," University of Iowa Working Paper No. 91-16, August.

Tamura, R. (1991b) "From Decay to Growth: A Dynamic Equilibrium Model of Income Distribution," University of Iowa mimeo, November.

Tamura, R. (1993) "Regional Economies and Market Integration," University of Iowa mimeo, January.

Tiebout (1956) "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64, 416-424.

Wilson, W. (1987) *"The Truly Disadvantaged,"* University of Chicago Press, Chicago.

Wheaton, W. (1993) "Land Capitalization, Tiebout Mobility and the Role of Zoning Regulations," *Journal of Urban Economics*, forthcoming.

</ref_section>