CHAPTER 1

Sorting, Education, and Inequality

Raquel Fernández

1. INTRODUCTION

Individuals sort in a variety of fashions. The workplace, the school of one's child, the choice of neighborhood in which to reside, and the selection of a spouse are all important arenas in which a choice of peers and access to particular goods and networks is explicitly or implicitly made. The aim of this chapter is to review the subset of the literature in the rapidly growing field of education and inequality that is primarily concerned with how individuals sort and the consequences of this for the accumulation of human capital, equity, efficiency, and welfare.

At first blush, sorting may seem like a rather strange lens through which to examine education. After all, this field has been primarily concerned with examining issues such as the returns to education, the nature of the education production function, or, at a more macro level, the relationship between education and per capita output growth.[1] A bit more thought, though, quickly reveals that sorting is an integral component of these questions. With whom one goes to school or works, who one's neighbors are, and who is a member of one's household are all likely to be important ingredients in determining both the resources devoted to and the returns to human capital accumulation.

It is interesting to note that in all these spheres there is at least some evidence indicating that sorting is increasing in the United States. Jargowsky (1996), for example, examines the changing pattern of residential segregation in the United States over the past few decades. He finds that although racial and ethnic segregation has stayed fairly constant (with some small decline in recent years), segregation by income has increased (for whites, blacks, and Hispanics) in all U.S. metropolitan areas from 1970 to 1990. This increased economic segregation, and the fact that schools increasingly track students by ability, suggests that there is likely to be increased sorting at the school or classroom level by

[1] For a survey of the education production function literature, see Hanushek (1986); for returns to education, see, for example, Heckman, Layne-Farrar, and Todd (1996); for education and growth, see, for example, Benhabib and Spiegel (1994).

income and ability. Kremer and Maskin (1996) find evidence for the United States, Britain, and France that there is increased sorting of workers into firms, with some high-tech firms (e.g., silicon valley firms) employing predominantly high-skilled workers and low-tech firms (e.g., the fast-food industry) employing predominantly low-skilled workers. Lastly, there is also some indication of greater sorting at the level of household partner (or "marital" sorting). Although the correlation between spousal partners in terms of years of education has not changed much over the past few decades (see Kremer, 1997), the conditional probability of some sociological barriers being crossed – for example, the probability that an individual with only a high-school education will match with another with a college education – has decreased, indicating greater household sorting (see Mare, 1991).

This chapter examines some of the literature that deals with the intersection of sorting, education, and inequality. This review is not meant to be exhaustive but to give a flavor of some of the advances in the theory and quantitative evidence. Furthermore, it should be noted that there is no overarching theoretical framework in this field. Rather, different models are interesting because of how they illuminate some of the particular interactions among these variables and others – for example, the role of politics, the interaction between private and public schools, or the efficacy of different mechanisms (e.g., markets vs. tournaments) in solving assignment problems. Thus, rather than sketch the contribution of each paper, I have chosen to discuss a few models in depth. Furthermore, as a primary concern in this area is the magnitude of different effects, wherever possible I focus on the contributions that have attempted to evaluate these.

The organization of the chapter is as follows. I begin with the topic of residential sorting. Local schooling is prevalent in most of the world. This policy easily leads to residential sorting and may have important implications for education and inequality, particularly in countries such as the United States in which the funding of education is also largely at the local level. I also use this section to review the theory of sorting. Next, I turn to examining sorting at the school level. The papers here are different, as they are primarily concerned with the interaction of public and private schools and with the properties of different mechanisms. Lastly I turn to recent work on household sorting and its consequences for education and inequality.

2. SORTING INTO NEIGHBORHOODS

Neighborhoods do not tend to be representative samples of the population as a whole. Why is this? Sorting into neighborhoods may occur because of preferences for amenities associated with a particular neighborhood (say, parks), because of some individuals' desire to live with some types of people or not to live with some others (say, ethnic groups who wish to live together in order to preserve their culture, or who end up doing so as a result of discrimination), and in response to economic incentives. This chapter is primarily concerned with the latter, and in particular with the endogenous sorting that occurs in response to economic incentives that arise as a result of education policies.

Primary and secondary education is a good that is provided locally. In industrialized countries, the overwhelming majority of children attend public schools (in the United States, this number was a bit over 91 percent in 1996; similar percentages exist in other countries).[2] Typically, children are required to live in a school's district to attend school there, making a neighborhood's school quality a primary concern of families in deciding where to reside. Furthermore, in most countries at least some school funding (usually that used to increase spending above some minimum) is provided locally; this is particularly true in the United States, where only 6.6 percent of funding is at the federal level, 48 percent is at the state level, and 42 percent is at the local level.[3]

Does it matter that education is provided at the local level? How does local provision of education affect the accumulation of human capital, its distribution, and efficiency in general? What are the dynamic consequences of local provision? How do other systems of financing and providing education compare? These are some of the questions explored in this section. I start with a brief overview of the economics of sorting, much of which carries through to the other sections.

2.1. Multicommunity Models: The Economics of Sorting

Characterizing equilibrium in models in which heterogeneous individuals can choose among a given number of potential residences, and in which these choices in aggregate affect the attributes of the community, is generally a difficult task. Since Westhoff (1977), economists working with these often called "multicommunity models" have tended to impose a single-crossing condition on preferences in order to obtain and characterize equilibria in which individuals either partially or completely separate out by type.[4] As discussed in further detail in the paragraphs that follow, the single-crossing condition also has two other very useful implications. First, it guarantees the existence of a majority voting equilibrium. Second, in many models it allows one to get rid of "trivial" equilibria (e.g., those in which all communities are identical) when a local stability condition is employed.

A typical multicommunity model consists of a given number of communities, each associated with a bundle (q, p). These bundles consist of a good, or input that is provided in some quality or quantity q *at the community level* and of a community level price p of some (usually other) good or service. The latter

[2] Digest of Education Statistics (1999).

[3] The remaining percentages come from other miscellaneous sources. These figures are for 1996–1997 (Digest of Education Statistics, 1999).

[4] In games of asymmetric information (e.g., signaling models and insurance provision), the assumption of single-crossing indifference curves is used to obtain either partial or completely separating equilibria.

can simply be a price associated with residing in the neighborhood, such as a local property tax. Thus, we can assess the indirect utility of an individual from these residing in a given community as $V(q, p; y)$, where y is an attribute of the individual such as income, ability, parental human capital, wealth, or taste. We assume throughout that q is "good" in the sense that $V_q > 0$, whereas $V_p < 0$.

Individuals choose a community in which to reside. In these models, equilibria in which individuals sort into communities along their characteristic y are obtained by requiring the slope of indifference curves in (q, p) space,

$$\left.\frac{dp}{dq}\right|_{v=\bar{v}} = -\frac{V_q}{V_p}, \tag{2.1}$$

to be everywhere increasing (or decreasing) in y. This implies that indifference curves cross only once and that where they do, if (2.1) is increasing in y, then the slope of the curve of an individual with a higher y is greater than one with a lower y; the opposite is true if (2.1) is decreasing in y.

The assumption of a slope that increases (decreases) in y ensures that if an individual with y_i prefers the bundle (q_j, p_j) offered by community j to some other bundle (q_k, p_k) offered by community k, and $p_j > p_k$, then the same preference ordering over these bundles is shared by all individuals with $y > y_i$ $(y < y_i)$. Alternatively, if the individual with y_i prefers (q_k, p_k), then community k will also be preferred to community j by all individuals with $y < y_i$ $(y > y_i)$.

Either an increasing or a decreasing slope can be used to obtain separation.[5] Henceforth, unless explicitly stated otherwise, I assume that (2.1) is increasing in y; that is,

$$\frac{\partial\left(\left.\frac{dp}{dq}\right|_{v=\bar{v}}\right)}{\partial y} = -\frac{V_{qy}V_p - V_{py}V_y}{V_p^2} > 0. \tag{2.2}$$

We shall refer to equilibria in which there is (at least some) separation by characteristic as sorting or stratification.

Condition (2.2) is very powerful. Independent of the magnitude of the expression, the fact that it is positive implies that individuals have an incentive to sort. As we discuss in the next section, this will be problematic for efficiency because it implies that even very small private incentives to sort will lead to a stratified equilibrium, independent of the overall social costs (which may be large) from doing so.

There are many economic situations in which condition (2.2) arises naturally. Suppose, for example, that q is the quality of education, and that this is determined by either a lump sum or a proportional tax p on income. If individuals

[5] Note that although either assumption can be used to obtain separation, the economic implications are very different. If increasing, then in a stratified equilibrium, higher-y individuals would obtain a higher (q, p). If decreasing, the high (q, p) bundle would be obtained by lower-y individuals.

are, for example, heterogeneous in income (so y denotes the income of the individual), then this condition would imply that higher-income individuals are willing to pay more (either in levels or as a proportion of their income, depending on the definition of p) to obtain a greater quality of education. This can then result in an equilibrium stratified along the dimension of income. Alternatively, if the quality of education is determined by the mean ability of individuals in the community school, p is the price of housing in the community, and individuals are heterogeneous in ability y, then (2.2) will be met if higher-ability individuals are willing to pay a higher price of housing to obtain higher-quality (mean ability) schooling, allowing the possibility of a stratified equilibrium along the ability dimension.

It is important to note, given the centrality of borrowing constraints in the human capital literature, that differential willingness to pay a given price is not the only criterion that determines whether sorting occurs.[6] Suppose, for example, that individuals are unable to borrow against future human capital or, less restrictively, that individuals with lower income, or lower wealth, or whose parents have a lower education level, face a higher cost of borrowing. Then even in models in which there is no other incentive to sort (e.g., in which the return to human capital is not increasing in parental assets or, more generally, in which V_q is not a function of y), there will nonetheless be an incentive to sort if the cost of residing in communities with higher qs (i.e., the effective p that individuals face) is decreasing in y. So, for example, if individuals with fewer assets face a higher effective cost of borrowing (they are charged higher rates of interest by banks), then they will be outbid by wealthier individuals for housing in communities with a higher q.

In many variants of multicommunity models, not only does (2.2) give rise to stratified equilibria, but it also implies that all locally stable equilibria must be stratified.[7] In particular, the equilibrium in which all communities offer the same bundle, and thus each contains a representative slice of the population, is locally unstable.[8]

There are many local stability concepts that can be imposed in multicommunity models. A particularly simple one is to define local stability as the property that the relocation of a small mass of individuals from one community to another implies that under the new configuration of (q, p) in these communities, the relocated individuals would prefer to reside in their original community. More rigorously, an equilibrium is locally stable if there exists an $\varepsilon > 0$, such that, for all possible combinations of measure δ $(0 < \delta \leq \varepsilon)$ of individuals $y_i \in \Lambda_j^*$ (where Λ_j^* is the set of individuals that in equilibrium reside in community j),

[6] For human capital models in which imperfections in credit markets play a central role, see Fernández and Rogerson (1998), Galor and Zeira (1993), Ljungqvist (1993), and Loury (1981), among others.

[7] In many settings, this gives rise to a unique locally stable equilibrium.

[8] Note that this zero sorting configuration is always an equilibrium in multicommunity models, as no single individual has an incentive to move.

a switch in residence from community j to k implies

$$V(q_k(\delta), p_k(\delta), y) \leq V(q_j(\delta), p_j(\delta), y), \quad \forall y \in \Lambda_{jk}, \forall j, k, \tag{2.3}$$

where $(q_l(\delta), p_l(\delta))$ are the new bundles of (q, p) that result in community $l = j, k$. Thus, condition (2.3) requires that, for all individuals who switch residence (the set Λ_{jk}), at the new bundles they should still prefer community j. This condition is required to hold for all community pairs considered.[9]

To see why the equilibrium with no sorting is rarely locally stable, consider, for example, the relocation of a small mass of high-y individuals from community j to k. In models in which the provision of the local good is decided by majority vote, this will tend to make the new community more attractive to the movers (and the old one less attractive), because the median voter in community k will now have preferences closer to those of the high-y individuals whereas the opposite will be true in community j. In models in which q is an increasing function of the mean of y (or an increasing function of an increasing function of the mean of y), such as when q is spending per student or the average of the human capital or ability of parents or students, then again this move will make community k more attractive than community j for the high-y movers. Thus, in all these cases, the no-sorting equilibrium will be unstable.

In several variants of multicommunity models, existence of an equilibrium (other than the unstable one with zero sorting) is not guaranteed.[10] For example, in a model in which the community bundle is decided upon by majority vote and voters take community composition as given, a locally stable equilibrium may fail to exist. The reason for this is that although there will exist (often infinite) sequences of community bundles that sort individuals into communities, majority vote need not generate any of these sequences. Introducing a good (e.g., housing) whose supply is fixed at the local level (so that the entire adjustment is in prices), though, will typically give rise to existence.[11]

Condition (2.2) also has an extremely useful implication for the political economy aspect of multicommunity models. Suppose that p and q are functions of some other variable t to be decided upon by majority vote by the population in the community (say, a local tax rate). They may also be functions of the characteristics of the (endogenous) population in the community. An implication of (2.2) is that independent of whether p and q are "nicely" behaved functions of t, the equilibrium outcome of majority vote over t will be the value preferred by the individual whose y is median in the community.

The proof of this is very simple. Consider the (feasible) bundle $(\tilde{q}, \tilde{p})$ preferred by the median-y individual in the community, henceforth denoted $\tilde{y}$. An

[9] See, for example, Fernández and Rogerson (1996). If communities have only a fixed number of slots for individuals as in models in which the quantity of housing is held fixed, then this definition must be amended to include the relocation of a corresponding mass of individuals from community k to j.

[10] See Westhoff (1977) and Rose-Ackerman (1979).

[11] See, for example, Nechyba (1997).

implication of (2.2) is that any feasible (q, p) bundle that is greater than $(\tilde{q}, \tilde{p})$ will be rejected by at least 50 percent of the residents in favor of $(\tilde{q}, \tilde{p})$, in particular by all those whose y is smaller than $\tilde{y}$. On the other hand, any feasible bundle with a (q, p) lower than $(\tilde{q}, \tilde{p})$ will also be rejected by 50 percent of the residents, namely all those with $y > \tilde{y}$. Thus, the bundle preferred by $\tilde{y}$ will be chosen by majority vote.[12]

It is also important to note that even in the absence of a single-crossing condition, to the extent that education is funded in a manner that implies redistribution at the local level, wealthier individuals will have an incentive to move away from less wealthy ones. This is by itself a powerful force that favors sorting but often requires a mechanism (e.g., zoning) to prevent poorer individuals from chasing richer individuals in order to enjoy both a higher q and a lower p.

For example, a system of local provision of education funded by a local property tax implicitly redistributes from those with more expensive housing to those with less expensive housing in the same neighborhood. The extent of redistribution, though, can be greatly minimized by zoning regulations that, for example, require minimum lot sizes.[13] This will raise the price of living with the wealthy and thus greatly diminish the amount of redistribution that occurs in equilibrium. In several models, to simplify matters, it will be assumed that mechanisms such as zoning ensure perfect sorting.

2.2. The Efficiency of Local Provision of Education

The simplest way to model the local provision of education is in a Tiebout model with (exogenously imposed) perfect sorting. In this model, individuals with different incomes y_i, but with identical preferences over consumption c and quality of education q, sort themselves into homogeneous communities. Each community maximizes the utility of its own representative individual subject to the individual or community budget constraint. Let us assume that the quality of education depends only on spending per student (i.e., the provision of education exhibits constant returns to scale and there are no peer effects). Then, perfect sorting is Pareto efficient. Note that this system is identical to a purely private system of education provision.

The model sketched in the previous paragraph often guides many people's intuition in the field of education. This is unfortunate, as it ignores many issues central to the provision of education. In particular, it ignores the fact that education is an *investment* that *benefits the child* and potentially affects the *welfare of others* as well. These are important considerations, as the fact that education is primarily an investment rather than a consumption good implies that borrowing constraints may have significant dynamic consequences; the fact that

12 See Westhoff (1977) and Epple and Romer (1991). Also see Gans and Smart (1996) for a more general ordinal version of single crossing and existence of majority vote.

13 See Fernández and Rogerson (1997b) for an analysis that endogenizes zoning, sorting, and the provision of education.

education primarily affects the child's (rather than the parents') welfare raises the possibility that parents may not be making the best decisions for the child. Furthermore, the potential externalities of an agent's education raise the usual problems for Pareto optimality.

In the paragraphs that follow, I explore some departures from the assumptions in the basic Tiebout framework and discuss how they lead to inefficiency of the stratified equilibrium. This makes clear a simple pervasive problem associated with sorting, namely that utility-maximizing individuals do not take into account the effect of their residence decisions on community variables. I start by discussing the simplest modification to the basic Tiebout model – reducing the number of communities relative to types.

Following Fernández and Rogerson (1996), consider an economy with a given number of communities $j = \{1, 2, \ldots, N\}$, each (endogenously) characterized by a proportional income tax rate t_j and a quality of education q_j equal to per pupil expenditure, that is, $q_j = t_j \mu_j$. Individuals who differ in income y_i, where $i \in I = \{1, 2, \ldots, \bar{I}\}$ (with $y_1 > y_2 > \cdots > y_{\bar{I}}$), simultaneously decide in which community, C_j, they wish to reside. Once that decision is made, communities choose tax rates by means of majority vote at the community level. Individuals then consume their after-tax income and obtain education.[14]

Assume for simplicity that individual preferences are characterized by the following separable specification,

$$u(c) + v(q), \tag{2.4}$$

so that sorting condition (2.2) is satisfied if $-\{[u''(c)c]/[u'(c)]\} > 1, \forall c$. We henceforth assume that the inequality is satisfied, ensuring that individuals with higher income are willing to suffer a higher tax rate for higher quality.[15]

Suppose that the number of communities is smaller than the number of income types.[16] In such a case the equilibrium will generally not be Pareto efficient. The clearest illustration of this can be given for the case in which individuals have preferences such that an increase in the mean income of the community *ceteris paribus* decreases the tax rate that any *given* individual would

[14] Very often, the literature in this field has implicitly adopted a sequencing such as the one outlined here. Making the order of moves explicit as in Fernández and Rogerson (1996) allows the properties of equilibrium (e.g., local stability) to be studied in a more rigorous fashion. It would also be of interest to examine properties of models in which communities act more strategically and take into account the effect of their tax rate on the community composition. There is no reason to believe that this modification would generate an efficient equilibrium, however.

[15] Most assumptions here are for simplicity only; for example, preferences need not be separable, and introducing housing and property taxation rather than income taxation would allow a sorting equilibrium to be characterized by higher-income communities having lower tax rates (but higher tax-inclusive prices) and higher q. We forgo the last option, as it simply complicates matters without contributing additional insights.

[16] Note that type here is synonymous with income level. Hence the assumption that there are fewer neighborhoods than types is a reasonable one to make.

like to impose. As the preferred tax rate of an individual is given by equating $u'(c)y_i$ to $v'(q)\mu_j$, this is ensured by assuming $-\{[v''(q)q]/[v'(q)]\} > 1$ (note that this is the parallel of the condition on u that generates sorting).[17]

As discussed previously, the result of majority vote at the community level is the preferred tax rate of the median-income individual in the community. There are a few things to note about the characteristics of equilibrium. First, in equilibrium, no community will be empty. If one were, then in any community that contained more than one income type, those with higher income would be made better off by moving to the empty community, imposing their preferred tax rate, and engaging in no redistribution. Second, in a locally stable equilibrium, communities cannot offer the same bundles and contain more than one type of individual (as a small measure of those with higher income could move to one of the communities, increase mean income there, and end up with the same or a higher-income median voter who has preferences closer to theirs). Lastly, if communities have different qualities of education (as they must if the communities are heterogeneous), then a community with a strictly higher q than another must also have a strictly higher t (otherwise no individual would choose to reside in the lower-quality–higher-tax community).

In the economic environment described herein, all locally stable equilibria must be stratified; that is, individuals will sort into communities by income. In such equilibria, communities can be ranked by the quality of education they offer, their income tax rate, and the income of the individuals that belong to them. Thus, all stable equilibria can be characterized by a ranking of communities such that $\forall j$, $q_j > q_{j+1}$, $t_j > t_{j+1}$, and $\min y_i \in C_j \geq \max y_i \in C_{j+1}$.

To facilitate the illustration of inefficiency, assume for simplicity that there are only two communities, $j = 1, 2$, and $I > 2$ types of individuals.[18] A stratified equilibrium will have all individuals with income strictly greater than some level y_b living in C_1 and those with income strictly lower than y_b living in C_2 with $q_2 > q_1$ and $t_2 > t_1$.

Suppose that in equilibrium individuals with income y_b live in both communities. It is easy to graph the utility

$$W_b^j \equiv u(y_b(1 - t_j)) + v(t_j\mu_j) \tag{2.5}$$

of these "boundary" individuals as a function of the community in which they reside and as a function of the fraction ρ_b of these individuals that reside in C_1. Let ρ_b^* denote the equilibrium value of the boundary individuals residing in C_1. Note that a decrease in ρ_b from its equilibrium value that does not alter the identity of the median voter in either community will make individuals with income y_b better off in both communities, as mean incomes will rise, qualities of education increase, and tax rates fall in both communities. Thus,

[17] This assumption implies that an increase in the mean income of the community that does not change the identity of the median voter will result in a higher q and a lower t, ensuring that all residents are made better off.

[18] See Fernández and Rogerson (1996) for a generalization of this argument to many communities.

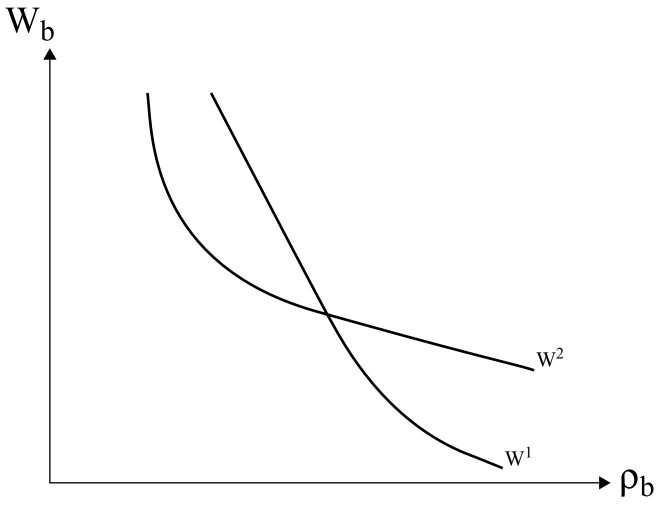

Figure 1.1. Equilibrium.

for this equilibrium to be locally stable, it must be that such a decrease makes y_b individuals even better off in C_1 relative to C_2, reversing the outward flow and reestablishing ρ_b^* as the equilibrium. Thus, as shown in Figure 1.1, the W_b^1 curve must cross the W_b^2 curve from above.[19]

This equilibrium is clearly inefficient. Consider a marginal subsidy of $s > 0$ to all individuals with income y_b who choose to reside in C_2.[20] Given that without a subsidy these individuals are indifferent between residing in either community, it follows that a subsidy will increase the attractiveness of C_2 relative to C_1. Consequently, some y_b individuals will move to C_2, thereby increasing mean income in both communities. For a small enough subsidy such that the identity of the median voter does not change in either community, the overall effect will be to decrease tax rates and increase the quality of education in both communities, thus making all individuals better off. Thus, it only remains to show that the subsidy can be financed in such a way to retain the Pareto-improving nature of this policy. A simple way to do so is by (marginally) taxing those y_b individuals who remain in C_1.[21] This tax will only further increase their outflow from C_1 to the point where they are once again indifferent between residing in both communities. As shown in Figure 1.1, the tax serves to further increase the utility of this income group (and consequently everyone else's). This last point suggests that a simpler way of producing the same

[19] Note that we are assuming for the range of ρ_b shown that neither of the communities' median voters are changing.

[20] If income is unobservable, then a small subsidy to all individuals who reside in C_2 would have to be paid.

[21] Again, if income is not observable, it is possible to preserve the Pareto-improving nature of this policy by (marginally) taxing all C_1 residents.