Inequality and Segregation*

Rajiv Sethi† Rohini Somanathan‡

February 11, 2002

Abstract

This paper explores the manner in which race and income interact to determine patterns of residential location in metropolitan areas. We use a framework in which individuals care about both the level of affluence and the racial composition of their communities, and in which there are differences in income both within and between groups. Three main findings emerge. First, under certain conditions there exist stable equilibria in which, *conditional on income*, black households experience lower neighborhood quality relative to whites. Second, extreme levels of segregation can be stable when racial income disparities are either sufficiently large or sufficiently small, but unstable in some intermediate range. Third, there exist multiple stable equilibria with very different levels of segregation when racial income disparities are sufficiently small. These results hold even when preferences are pro-integrationist, in the sense that racially mixed neighborhoods within a certain range are strictly preferred by all households to homogenous neighborhoods of either type.

*We thank Dilip Abreu, JP Benoit, Duncan Foley, Malgosia Madajewicz, Lalith Munasinghe, Dan O'Flaherty, Efe Ok, Debraj Ray, David Weiman, and seminar participants at Columbia, Michigan, and NYU for helpful comments. An earlier version of this paper was circulated under the title "Race, Income and Neighborhood Sorting."

†Department of Economics, Barnard College, Columbia University (rs328@columbia.edu).

‡Department of Economics, University of Michigan, and Planning Unit, Indian Statistical Institute - Delhi (rohinis@umich.edu).

1 Introduction

Several decades have elapsed since the landmark Civil Rights Act of 1964 outlawed discrimination in employment and public education, and the 1968 Fair Housing Act extended these protections to the sale and rent of housing. Over this period racial disparities in educational attainment and household income have narrowed, a significant population of middle class African Americans has emerged, and the attitudes of white Americans towards integrated schools and neighborhoods have softened considerably.[1]

In comparison with changes in laws, attitudes and the racial composition of the middle class, changes in black-white residential segregation over the past thirty years have been gradual and uneven. Many areas in which black households constitute a small percentage of the metropolitan population have witnessed greater integration, but segregation remains a striking feature of the urban landscape in several cities with large black populations. Moreover, the levels of segregation experienced by black households are uniformly high across all income categories, and the relative segregation of different cities has remained remarkably stable for over a century.[2]

The persistence of segregation has generally been attributed to some combination of racial income disparities, discrimination in real estate and credit markets, and preferences over neighborhood racial composition. Implicit in this characterization is the presumption that declining economic inequality should, *ceteris paribus*, result in lower segregation. This would certainly be the case if sorting across neighborhoods were based on income alone. When individuals care also about the racial composition of their communities, however, the relationship between inequality and segregation is more complex and depends in subtle ways on both intraracial and interracial disparities in income. It is with this relationship that the present paper is concerned.

We analyze a model in which incomes vary both within and between groups, and indi-

[1] Approximately one-half of black Americans now live in middle or upper income households as compared with about one-fifth in 1960. The black-white gap in high school completion rates for 25-29 year olds dropped from 20 percentage points in 1967 to 7 points in 1996. Median black household income rose by 41% between 1967 and 1999 while the median white household income rose by 24% (U.S. Census Bureau, 2000, Council of Economic Advisors, 1998). In 1963 only 39% of white respondents disagreed with the statement that whites had a right to keep blacks out of their neighborhood; by 1996 this had risen to 86% (Schuman et al., 1997). Additional evidence on attitudes is discussed in Section 2 below.

[2] See Denton and Massey (1988), Massey and Denton (1993), Farley and Frey (1994), and Cutler et al. (1999) for evidence and interpretations, and Glaeser and Vigdor (2001) and Lewis Mumford Center (2001) for an analysis of the most recent data.

viduals care about both the level of affluence and the racial composition of the communities in which they reside. This concern with racial composition may be pro-integrationist, in that households prefer some degree of mixing to homogenous neighborhoods of either type. Individuals are able to locate in any neighborhood, provided that they are willing and able to outbid others to do so.[3] We focus on equilibrium allocations that are *stable* in the sense that small perturbations in the neighborhood of the equilibrium are self-correcting under the dynamics of decentralized neighborhood choice.

We obtain three main results. First, under certain conditions, there exist stable equilibria at which, *conditional on income*, black households experience lower neighborhood quality than white households. Such equilibria arise even when preferences over neighborhood racial composition are strongly pro-integrationist. This result depends crucially on the hypothesis that neighborhood racial composition matters to households, and could not arise if location decisions were based on income alone.[4]

Second, extreme levels of segregation can be stable if racial income disparities are either very great or very small, but unstable in some intermediate range. When income disparities are large, segregation results from black households not being rich enough to afford the high rents of more affluent white neighborhoods, and when income disparities are negligible, it results from the fact that high income black households have less to gain in terms of improved neighborhood quality by moving to a white neighborhood. When racial income disparities are neither too large nor too small, on the other hand, segregation may be unstable so that the only stable allocations are integrated. A narrowing of racial income disparities is therefore consistent with *increasing* segregation and, from a cross-sectional perspective, one ought not to expect cities with the smallest racial income disparities to be the ones with the lowest levels of segregation.

Finally we show that when racial income disparities are sufficiently small, multiple equilibria exist, and these equilibria can differ dramatically in their corresponding levels of segregation. The existence of multiple equilibria suggests that although integration may become viable as racial income disparities lessen, history may trap a city in a segregated equilibrium.

[3] This may not be possible in practice due to racial steering by real estate agents or discrimination in mortgage lending markets (Yinger, 1995). We abstract from such overt discrimination because its effects on segregation are reasonably well understood, and because doing so allows us to better focus on the questions at hand.

[4] The fact that white households experience higher neighborhood mean income than equally affluent black households has implications for the transmission of racial inequality across generations (Loury, 1977). These intergenerational effects are not explored in the present paper but clearly constitute an important extension.

This is where social policy may be most effective: temporary incentives for integration may give rise to permanent effects. Integration comes at the cost of higher stratification by income, however, so integrationist policies need not be unambiguously welfare enhancing even when preferences are strongly pro-integrationist.

Our work is closely related to two literatures which deal with the decentralized dynamics of neighborhood choice. The idea that extreme levels of segregation can arise under a broad range of preferences over neighborhood composition was developed in seminal work by Schelling (1971, 1972). This analysis contains the important insight that even when all individuals prefer integrated neighborhoods to segregated ones, integration may be unsustainable in that a few random shocks can tip the system to a segregated equilibrium. It is difficult, therefore, to deduce anything about individual preferences from aggregate patterns of residential location.[5] While Schelling's analysis neglects the role of prices in rationing housing demand, broadly similar conclusions hold in models that take full account of adjustments in rents (Yinger, 1976, Schnare and McRae, 1978, Kern, 1981). This literature neglects the fact that individuals consider both race and income when making location choices, and that forces acting to produce stratification by income can substantially mitigate the amount of racial segregation that results. While extreme levels of segregation are consistent with pro-integrationist preferences in our model, it is also the case that, under certain circumstances, stable equilibria can entail greater *integration* than any individual, black or white, considers ideal.

There is also an extensive literature on neighborhood sorting when individuals differ with respect to their incomes and sort themselves across jurisdictions on the basis of neighborhood characteristics such as local taxation, redistribution, public education, or peer-effects (De Bartolome, 1990, Epple and Romer, 1991, Benabou, 1992, Fernandez and Rogerson, 1996, Durlauf, 1996, Epple and Platt, 1998). Stratification by income occurs in many such models. What is missing from this body of work is the possibility that individuals care about certain intrinsic characteristics of those with whom they share their neighborhoods, that such preferences are themselves related to group membership. When there is inequality both within and between groups, adding these components to the analysis yields significant new insights that appear neither in the segregation literature descended from Schelling, nor

[5] "People who have to choose between polarized extremes ... will often choose in a way that reinforces the polarization. Doing so is no evidence that they prefer segregation, only that, if segregation exists and they have to choose between exclusive association, people elect like rather than unlike environments." (Schelling, 1978, p.146).

in the literature on neighborhood sorting in the Tiebout tradition.

The paper is organized as follows. Section 2 provides some discussion and justification for our key assumption that individuals care about both the racial composition and the level of affluence in their communities. The model is developed in Section 3, and its equilibrium properties characterized in Sections 4 and 5. Section 6 examines the relationship between racial income disparities and residential segregation, and Section 7 concludes. All proofs are in an Appendix.

2 Preferences

Extensive survey evidence on the racial attitudes of Americans has been collected for more than half a century (Schuman et al., 1997). Several studies have specifically attempted to ascertain the preferences of respondents over neighborhood racial composition (Farley et. al., 1978, 1993, Bobo et al., 1986). The best recent evidence comes from a 'Multi-City Study of Urban Inequality' funded jointly by the Ford Foundation and the Russell Sage Foundation. Subjects drawn from the Los Angeles and Boston metropolitan areas were asked to construct an "ideal neighborhood that had the ethnic and racial mix" that the respondent "personally would feel most comfortable in". They did so by examining a card depicting three rows of five houses each, imagining their own house to be at the center of the middle row, and assigning to each of remaining houses an ethnic/racial category using the letters A (Asian), B (Black), W (White) and H (Hispanic). The study found evidence that "all groups prefer both substantial numbers of co-ethnic neighbors and considerable integration" (Zubrinsky Charles, 2001, p.257). On average, the ideal neighborhood consisted of a plurality of the respondent's own type (ranging from 40% for black respondents to 52% for whites) together with significant representation from other groups. Only 2.5% of blacks to 11.1% of whites considered homogeneous neighborhoods populated only with their own type to be ideal. Overall, this reflects a clear desire for some degree of integration on the part of all groups, with a bias towards members of one's own group. This is consistent with prior studies of attitudes towards racial composition and motivates the specification used in this paper.

Why might individuals care about the racial composition of their neighborhoods? Farley et al. (1994) trace white attitudes to negative racial stereotypes, and black attitudes to anticipated hostility from whites. Ellen (2000) argues that white households hold an exaggerated view (relative to black households) of the association between changes in racial composition and structural decline in neighborhood quality. Whites are consequently less willing than blacks to settle in neighborhoods which have recently experienced increases in the share of the black residents. O'Flaherty (1999) has argued that interracial transactions of many kinds are rendered difficult because the signals blacks and whites send each other through their actions and words "are garbled by stereotypes and the possibility of animosity." The fact that communication is easier and less ambiguous when it does not cross racial lines could account for a desire to live with a substantial number of co-ethnics. Signals also play a key role in the search-theoretic model of Lundberg and Startz (1998), where signals from members of one's own group are interpreted with less noise than signals from others. Again this can lead endogenously to a desire to associate primarily with co-ethnics. While we take preferences over neighborhood racial composition to be exogenously given, our specification is consistent with these interpretations. In addition, we allow for the possibility that there may be a preference for some degree of integration on the part of both blacks and whites, as suggested by the survey evidence.

In addition to a concern about neighborhood racial composition, we assume that individuals also care about the level of affluence in their communities. There are a number of reasons why this might be the case. The quality of public schools is liable to be better in more affluent neighborhoods even if government per-pupil expenditures are uniform across the city. This is the case because voluntary contributions to parent-teacher associations increase with income, and because human capital transfers that occur in the home have spillover effects in school. The presence of positive role models (and the absence of negative ones) is correlated with the degree of affluence of a community. Living an a more affluent community provides entry into social networks which can be lucrative. And if the external upkeep of one's residence is a normal good with positive external effects, more affluent communities will be more desirable. Each of these effects have been discussed extensively in the literature (Bond and Coulson, 1984, De Bartolome, 1990, and Benabou, 1992). Although the desire to live in a more affluent community can be endogenously derived on the basis of any of the above concerns, it is treated here as a primitive of the model.

3 The Model

Consider a city with a continuum of households represented by the interval $[0, 1]$. Households differ along two dimensions, income and race. There are two races, black and white, and the share of black households in the city is denoted $\beta \leq \frac{1}{2}$. Within each racial group the income distribution is represented by absolutely continuous distribution functions $F^b(y)$ and

$F^w(y)$, with $f^b(y)$ and $f^w(y)$ denoting the corresponding densities. The supports of the two income distributions are given by the intervals $[y^b_{\min}, y^b_{\max}]$ and $[y^w_{\min}, y^w_{\max}]$. It is assumed that $y^b_{\min} \leq y^w_{\min} < y^b_{\max} \leq y^w_{\max}$, and for any $y \in (y^b_{\min}, y^w_{\max})$, $F^b(y) > F^w(y)$. Taken together, these assumptions imply that whites are wealthier than blacks as a group, although the wealthiest black households are better off than the poorest white ones.

The city is divided into two disjoint neighborhoods of equal size.[6] Any subset $A \subset [0,1]$ with measure one-half represents an allocation of households across neighborhoods, with the interpretation that $i \in A$ implies that household i resides in neighborhood 1, while the remaining households are in neighborhood 2. Any allocation of households across neighborhoods will imply both a racial composition and a distribution of income within each neighborhood. Let $\bar{y}_j(A)$ denote the mean income in neighborhood $j \in \{1,2\}$, $\beta_j(A) \in [0, 2\beta]$ the share of neighborhood j's population that is black, and $\omega_j(A) = 1 - \beta_j(A)$ the share of neighborhood j's population that is white.

Housing units are identical, and rents are accordingly uniform within each neighborhood. We normalize the rent in neighborhood 1 to equal zero and let ρ be the (possibly negative) rent in neighborhood 2. All income not spent on rent is used for private consumption. Apart from their private consumption, individuals care about the general affluence and racial composition of their communities. Neighborhoods with higher mean incomes are more desirable than those with lower mean incomes for all members of the population. Additionally, black and white households differ systematically with regard to their preferences over neighborhood racial composition. We shall assume for simplicity that the preferences of blacks and whites are symmetric in a sense to be made clear below. We do not assume, however, that preferences are monotonic in neighborhood racial composition. In particular, we allow for the possibility that households strictly prefer a wide range of integrated neighborhoods to segregated ones, and that being part of a sizeable minority may be more attractive than being part of an overwhelming majority.

Preferences are represented by the following simple, separable utility function

$$U(c, \bar{y}, r) = u(c) + \bar{y} + v(r),$$

where c is private consumption, $\bar{y}$ is neighborhood mean income, $r \in \{\beta, \omega\}$ is the neighborhood population share of the individual's own race, $u(c)$ is continuous, strictly increasing

[6] Although the analysis in this paper is focused on the two neighborhood case, the main results characterizing the relationship between racial income disparities and equilibrium segregation (Propositions 3 and 4) can be appropriately modified to hold also in the case of an arbitrary number of neighborhoods.

and strictly concave, and $v(r)$ is given by

$$v(r) = r\,(1 - r + \eta)\,. \tag{1}$$

The parameter $\eta \in [0, 1]$ measures the degree to which residence with co-ethnics is desired. When $\eta = 0$ each individual's ideal neighborhood racial composition consists of equal shares of blacks and whites. More generally, the ideal racial composition for an individual is to have a share $\frac{1}{2}(1 + \eta)$ of her own type in the neighborhood. Larger values of η therefore correspond to a greater bias towards one's own group. Except in the extreme case $\eta = 1$, such preferences are nonmonotonic: all individuals prefer some degree of integration to complete segregation. For any value of $\eta < \frac{1}{2}$, the range of neighborhood compositions that are strictly preferred to complete segregation includes allocations in which the individual is in a minority.

Equilibrium in this model is an allocation A of households across neighborhoods and a neighborhood 2 rent ρ such that no household prefers a neighborhood different from its own. In other words, equilibrium requires that for any household $i \in [0, 1]$ with income y, it must be the case that

$$\begin{aligned} U(y, \bar{y}_1, r_1) &\geq U(y - \rho, \bar{y}_2, r_2) \quad \text{if } i \in A \\ U(y, \bar{y}_1, r_1) &\leq U(y - \rho, \bar{y}_2, r_2) \quad \text{if } i \notin A \end{aligned}$$

where $r = \beta$ for black households, $r = \omega$ for white households. We shall refer to an allocation in which each neighborhood contains members of both races as *integrated*, and all other allocations as *segregated*.

4 Intraracial Stratification

We say that an allocation is *intraracially stratified* if there exist threshold income levels $\tilde{y}^b$ and $\tilde{y}^w$ such that one neighborhood consists exclusively of all black households with income above $\tilde{y}^b$ together with all white households with income above $\tilde{y}^w$. The neighborhood with this property shall be referred to as the *upper-tail neighborhood*. The other (lower-tail) neighborhood then consists exclusively of all black households with income below $\tilde{y}^b$ together with all white households with income below $\tilde{y}^w$. Intra-racial stratification is consistent with complete segregation (if $F(\tilde{y}^b) = 0$ or 1), with pure stratification by income (if $\tilde{y}^b = \tilde{y}^w$), and a variety of other patterns of neighborhood sorting including equal neighborhood racial compositions and equal neighborhood mean incomes. Without loss of generality, we adopt the convention that at any intraracially stratified allocation, neighborhood 2 is the upper-tail neighborhood. For reasons discussed below, the lower-tail neighborhood could have greater mean income than the upper-tail neighborhood at some equilibrium allocations.

When an allocation is intraracially stratified, the mean incomes and racial compositions in each neighborhood can all be expressed as a function of the threshold income level for white households. Let $z = \tilde{y}^w$ denote the threshold income for whites. It must be the case that $z \in [z_{\min}, z_{\max}]$ where $z_{\min}$ is defined by the condition $(1-\beta)(1 - F^w(z_{\min})) = \frac{1}{2}$, and $z_{\max}$ by $(1-\beta) F^w(z_{\max}) = \frac{1}{2}$. When $z = z_{\min}$ the upper-tail neighborhood consists exclusively of white households, and when $z = z_{\max}$ the lower-tail neighborhood is exclusively white. Given any value of $z \in [z_{\min}, z_{\max}]$, there exists a unique $\tilde{y}^b \in [y^b_{\min}, y^b_{\max}]$ such that

$$\beta F^b(\tilde{y}^b) + (1-\beta) F^w(z) = \frac{1}{2}.$$

The threshold $\tilde{y}^b(z)$ identifies the unique level of black income such that the blacks with income above this threshold and whites with income above z together constitute half the population. Note that $\tilde{y}^b(z)$ is a continuous, strictly decreasing function on $[z_{\min}, z_{\max}]$.

At any intraracially stratified allocation z, the share of black households in neighborhood 1 is given by

$$\beta_1(z) = 2\beta F^b\left(\tilde{y}^b(z)\right),$$

where $\beta_1(z) \in [0, 2\beta]$ and $\beta_2(z) = 2\beta - \beta_1(z)$. Mean incomes in the two neighborhoods are

$$\begin{aligned}
\bar{y}_1(z) &= \frac{\beta_1(z)}{F^b\left(\tilde{y}^b(z)\right)} \int_{y^b_{\min}}^{\tilde{y}^b(z)} y f^b(y) dy + \frac{1-\beta_1(z)}{F^w(z)} \int_{y^w_{\min}}^{z} y f^w(y) dy \\
\bar{y}_2(z) &= \frac{\beta_2(z)}{1 - F^b(\tilde{y}^b(z))} \int_{\tilde{y}^b(z)}^{y^b_{\max}} y f^b(y) dy + \frac{1-\beta_2(z)}{1 - F^w(z)} \int_{z}^{y^w_{\max}} y f^w(y) dy.
\end{aligned}$$

Hence all neighborhood characteristics relevant to households are fully determined by the threshold white income z.

When $z = z_{\min}$, there is complete residential segregation by race, with the second (all-white) neighborhood having higher mean income. As z rises from this minimum value, the lowest income whites in the second neighborhood are replaced by the highest income blacks from the first, which leads to increasing income disparities across neighborhoods. The point at which neighborhood income disparities are greatest occurs when $\tilde{y}_b(z) = z$. This would be the outcome if sorting were based on income alone.[7] As z rises beyond this

[7] If the largest black income lies below the median white income and if the black share of the metropolitan population is sufficiently small, there may be no $z \in [z_{\min}, z_{\max}]$ such that $\tilde{y}^b(z) = z$. In this case sorting by income alone would give rise to a segregated allocation and the largest income difference between the two neighborhoods would occur when $z = z_{\min}$.

point, overall stratification by income begins to decline. If β is sufficiently large, at some point the two neighborhoods have identical mean incomes; beyond this the second (upper-tail) neighborhood has lower mean income since it consists of all but the poorest segments of the less affluent race together with a few of the wealthiest members of the more affluent race. Finally, when $z = z_{\max}$, the allocation is again segregated but with the most affluent whites sharing a neighborhood with the city's black population, while lower income whites live in a racially homogenous neighborhood.

We shall say that an allocation is *symmetric* if the two neighborhoods are identical with respect to both mean income and racial composition. It is easily shown that symmetric allocations exist and that any symmetric allocation will be an equilibrium when the neighborhoods have equal rents. Moreover:

Proposition 1 *Equilibrium allocations must be either segregated, symmetric, or intraracially stratified.*

Equilibrium allocations that are not intraracially stratified must therefore be either segregated or symmetric. Symmetric equilibria will be unstable under the dynamics of decentralized location choices. For instance, the movement of the wealthiest white households from one neighborhood to the other in exchange for poorer white households will result in disparities in mean income while preserving the equality of racial compositions. This would result in wealthier households (of both races) outbidding poorer ones to locate in the more affluent community. Hence stable integrated equilibria must be intraracially stratified.

Segregated equilibria which are *not* intraracially stratified can exist. Specifically, there can be an equilibrium in which the two neighborhoods have equal rents, all black households occupy the same neighborhood, and all white households are indifferent between the two neighborhoods. Since this allocation is not symmetric, the difference in mean income across neighborhoods must exactly compensate whites for differences in neighborhood racial composition. As in the case of symmetric allocations, such equilibria will also be unstable. A small movement of wealthier whites to one neighborhood in exchange for poorer whites will leave the racial compositions unaffected but tilt income disparities in such a way as to make whites strictly prefer the neighborhood which experiences the increase in mean income. Hence stable segregated equilibria must be intraracially stratified. Since stable integrated equilibria must also be intraracially stratified, we confine our attention in the remainder of this paper to intraracially stratified allocations. It is easily shown that $\rho \geq 0$ at any such equilibria:

the upper-tail neighborhood cannot have lower rent than the lower tail neighborhood, even if the latter has higher mean income.

The following two examples illustrate the range of equilibrium possibilities.

Example 1 *Suppose that the income distributions are uniform with support* $[0, 0.7]$ *for black households and* $[0.3, 1]$ *for white households,* $u(c) = \log c$, $\beta = 0.45$ *and* $\eta = 0.15$. *Then there are two equilibria:*

z^*	$\tilde{y}_b$	β_1	β_2	$\bar{y}_1$	$\bar{y}_2$	ρ
0.36	0.70	0.90	0.00	0.35	0.68	0.11
0.83	0.12	0.16	0.74	0.49	0.54	0.02

The first of these equilibria is segregated, with all black households in the first (lower-tail) neighborhood together with a few of the lowest income white households. Neighborhood 2 is more affluent and exclusively white. The marginal white household, despite being of lower income than the most affluent black household, is willing to outbid the latter to live in the second neighborhood. This is a reflection of black-white differences in preferences over neighborhood racial composition. The second equilibrium is unusual in that the second neighborhood is both more affluent and predominantly black. The fact that the most affluent whites are willing to pay the equilibrium rent to live there is a reflection of the fact that preferences are pro-integrationist. We show below, however, that this latter equilibrium is unstable in a well-defined sense.

Example 2 *Suppose that all specifications are as in Example 1, except that* $\eta = 0.08$. *Again there are two equilibria:*

z^*	$\tilde{y}_b$	β_1	β_2	$\bar{y}_1$	$\bar{y}_2$	ρ
0.46	0.58	0.75	0.15	0.31	0.72	0.15
0.86	0.09	0.12	0.78	0.52	0.51	0.01

Both equilibria in this example involve integration. As in the case of the segregated equilibrium of the previous example, the marginal white household, despite being of lower income than the marginal black household, is willing to outbid the latter to live in the more affluent upper-tail neighborhood. The second equilibrium starkly illustrates the impact of pro-integrationist preferences. The upper-tail neighborhood in this case is predominantly black and has *lower* income than the lower-tail neighborhood. Yet the wealthiest whites are willing to pay the equilibrium rent to live there. We show below that this equilibrium is unstable.

The next section introduces a simple notion of stability and shows that equilibria of the first type, in which the marginal white household is of lower income than the marginal black household, exist and are stable provided certain conditions are met.

5 Stable Equilibria

At any intraracially stratified allocation z, define the marginal bid-rent $\rho^w(z)$ as the maximum rent that a white household with income z is willing to pay to live in the upper-tail neighborhood. Similarly, define $\rho^b(z)$ as the maximum rent that a black household with income $\tilde{y}^b(z)$ is willing to pay to live in the upper-tail neighborhood. These functions are defined implicitly by the indifference conditions $F^w(z, \rho^w) = 0$ and $F^b(z, \rho^b) = 0$, where

$$\begin{aligned} F^w(z, \rho^w) &= U(z - \rho^w, \bar{y}_2(z), \omega_2(z)) - U(z, \bar{y}_1(z), \omega_1(z)), \\ F^b(z, \rho^b) &= U(\tilde{y}^b(z) - \rho^b, \bar{y}_2(z), \beta_2(z)) - U(\tilde{y}^b(z), \bar{y}_1(z), \beta_1(z)). \end{aligned}$$

There will always exist finite pair of marginal bid rents which satisfy these indifference conditions provided that $u(c)$ shows enough variation. We assume that this is indeed the case. Furthermore, this pair of marginal bid-rents is uniquely determined since $u(c)$ is strictly increasing.

Any integrated allocation at which the two marginal bid-rents $\rho^w(z)$ and $\rho^b(z)$ coincide and are positive is an equilibrium allocation with the equilibrium rent being equal to the common marginal bid-rents. This is the case because the marginal households (with income z and $y^b(z)$ respectively) are indifferent between the two neighborhoods, while all other households have a strict preference for the neighborhood to which they are allocated. Moreover, all integrated equilibria must be such that the marginal bid-rents coincide and are non-negative. At segregated equilibria the marginal bid-rents will generally differ, but it must be the case that all black households prefer the neighborhood in which they all reside. Hence if all black households live in the lower-tail neighborhood, we must have $\rho^w(z) \geq \rho^b(z)$, while the inequality is reversed if all black households live in the upper-tail neighborhood.

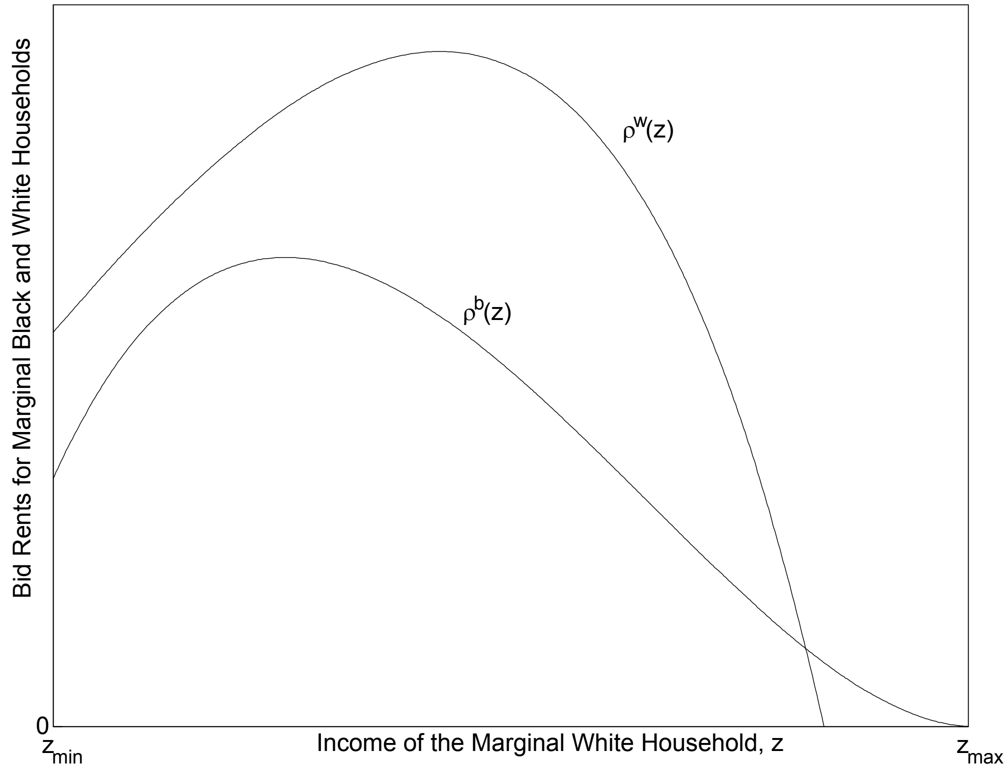

Figure 1. Marginal Bid Rent Curves: Stable Segregation

The marginal bid-rent functions for Example 1 are shown in Figure 1. As noted above, the integrated equilibrium is unusual in that the upper-tail neighborhood, despite being more affluent, is predominantly occupied by members of the less affluent race. In fact, it is easy to see that this equilibrium must be *unstable*: a small perturbation in the allocation of households would lead to cumulative divergence away from this allocation. For z slightly below the equilibrium value z^*, $\rho^w(z) > \rho^b(z)$. Hence the marginal white household is willing to pay more than the marginal black household to live in the upper-tail neighborhood. This would lead to an inflow of whites into the second neighborhood, further reducing z. Since there is no other point at which $\rho^w(z) = \rho^b(z)$, the only *stable* equilibrium involves complete segregation.

More formally, we say that an equilibrium allocation z^* is *stable* if there exists a neighborhood $N(z^*)$ such that, for all $z \in N(z^*) \cap [z_{\min}, z_{\max}]$ with $z \neq z^*$, $\rho^w(z) - \rho^b(z)$ has the same sign as $z - z^*$. If an equilibrium is not stable, it is *unstable*. This definition of stability implicitly assumes that when individuals relocate, they do so in a manner that maintains intraracial stratification: the marginal households are the first to move. Intuitively, an equilibrium allocation z^* is stable if sufficiently small perturbations of z away from z^* in either direction are self-correcting.

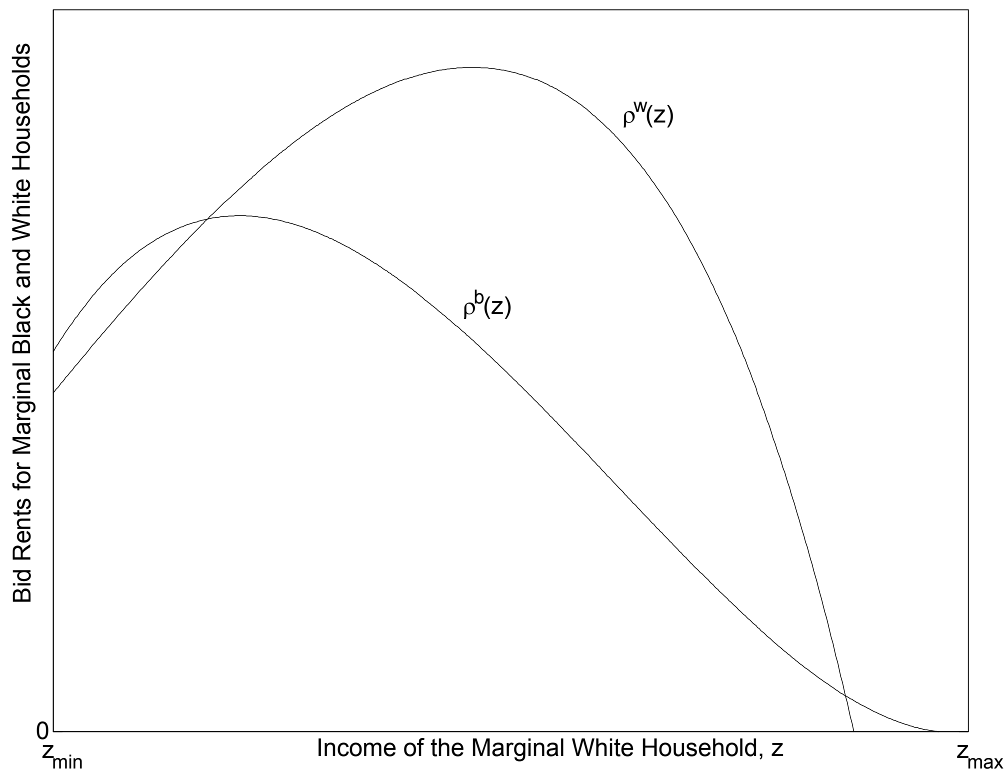

Figure 2. Marginal Bid Rent Curves: Stable Integration

In Figure 1, the only stable equilibrium involves segregation. The marginal bid-rent functions for Example 2 are shown in Figure 2. Here segregation is unstable since $\rho^b(z_{\min}) > \rho^w(z_{\min})$. The wealthiest black households are now willing to outbid the poorest white households to live in an all-white neighborhood, in order that they may benefit from the higher mean income. The unique stable equilibrium is now integrated, and there is a second integrated equilibrium which is unstable. Socioeconomic stratification in the stable equilibrium is higher than in the previous example but still falls short of the maximum possible stratification. Increasing tolerance leads to greater integration, as one might expect, but it also leads to greater disparities in mean income across neighborhoods. Furthermore the unique stable equilibrium is characterized by the property that the marginal black household is more affluent than the marginal white household ($\tilde{y}^b(z) > z$). The following result shows that a stable equilibrium of this kind must always exist provided the black share of the population is sufficiently large and a further condition is met.

Proposition 2 *Consider any $\eta > 0$ and suppose that $\beta > (1-\eta)/2$. Then, if $y^b_{\max} > z_{\min}$, there exists a stable equilibrium (z^*, ρ^*) such that $\tilde{y}^b(z^*) > z^*$.*

The condition $y^b_{\max} > z_{\min}$ is a weak one, requiring only that the most affluent black household in the city is better off than the lowest income white in the upper-tail neighborhood under complete segregation. Since $z_{\min}$ lies below the median income of white households, a sufficient condition for $y^b_{\max} > z_{\min}$ to hold for any β is that the highest black income exceed the median white income.

The equilibrium identified in Proposition 2 has the property that the wealthiest households in the lower income neighborhood will be black. In other words, there exists a range of incomes lying between z and $\tilde{y}^b(z)$ such that households falling within this range will be in the poorer neighborhood if and only if they are black. This group of households will also have higher levels of private consumption than white households with comparable incomes, since a smaller share of income is spent on housing. White and black households with the same income will therefore experience systematically different levels of neighborhood quality. Note that Proposition 2 holds no matter how close to perfectly pro-integrationist ($\eta = 0$) preferences happen to be, and no matter how much integration occurs in equilibrium: under the stated conditions there will always exist a stable equilibrium in which a set of households experience lower neighborhood quality *conditional on income* if they are black. In the presence of human capital externalities, the income of this group of black households will underpredict the future economic success of their children relative to the income of white households, an effect that could not occur under stratification alone. This is a sobering thought. Even in a world without overt discrimination, and one in which the desire for integration is strong, the advantage of being born to affluence may be magnified if one is also born to an affluent race.

Proposition 2 does not rule out stable equilibria in which $\tilde{y}^b(z) \leq z$, although the above examples clearly demonstrate that none may exist. Furthermore, it can be shown that at any integrated equilibrium with $\tilde{y}^b(z) \leq z$, the upper-tail neighborhood must have a larger share of black households than the lower-tail neighborhood, and that there exist at least as many unstable equilibria with this property as stable ones.

6 Income Disparities and Segregation

We turn next to the relationship between income disparities and the extent of segregation, and approach this question by allowing income distributions to depend on a scaling parameter $\alpha \in [0, 1]$. Let $F^b(y, \alpha)$ and $F^w(y, \alpha)$ represent these distributions. The corresponding supports are $[y^b_{\min}(\alpha), y^b_{\max}(\alpha)]$ and $[y^w_{\min}(\alpha), y^w_{\max}(\alpha)]$, assumed to be bounded and nondegenerate for all $\alpha \in [0, 1]$, and overlapping for all $\alpha \in (0, 1]$. It is further assumed that the distribution functions and their supports are continuous in α, and that higher values of α correspond to smaller racial income disparities. Specifically, $F^b(y, \alpha)$ is nonincreasing in α and $F^w(y, \alpha)$ is nondecreasing in α. Furthermore, $y^b_{\max}(0) = y^w_{\min}(0)$, and, for all y, $F^b(y, 1) = F^w(y, 1)$. The scaling parameter α represents racial income disparities, with $\alpha = 0$ corresponding to a completely hierarchical distribution of income and $\alpha = 1$ to identical income distributions. Racial disparities in the distribution of income can be tracked by looking at changes in α. The following result identifies conditions under which complete segregation is a stable equilibrium.

Proposition 3 *Consider any $\eta > 0$ and suppose that $\beta > (1 - \eta)/2$. Then (a) there exists $\alpha_l > 0$ such that if $\alpha < \alpha_l$, a stable, segregated equilibrium exists, and (b) there exist $\beta_h < \frac{1}{2}$ and $\alpha_h < 1$ such that if $\beta > \beta_h$ and $\alpha > \alpha_h$ both hold, a stable, segregated equilibrium exists.*

Proposition 3 states that complete segregation is stable whenever racial income disparities are either sufficiently large or sufficiently *small*. When racial income disparities are large, even allocations involving pure stratification by income are highly segregated, and preferences over neighborhood racial composition reinforce and exacerbate this effect. Hence the stability of complete segregation in this case is not surprising. The second and less obvious part of the result states that segregation will be stable in cities with significant black populations provided that racial income disparities are sufficiently small. This occurs because, when the two income distributions are virtually identical, complete segregation does not result in substantial income disparities across neighborhoods. This in turn implies that the benefit to wealthier black households from moving to higher income, predominantly white neighborhoods is small. Even a slight preference for all-black over all-white neighborhoods can overwhelm this effect and lead to stable patterns of extreme segregation. Consequently, the relationship between racial income disparities and the stability of segregated equilibria is nonmonotonic: segregation may be inconsistent with intermediate values of α while it is consistent with values of α lying at either extreme.

Part (b) of Proposition 3 need not hold when β is sufficiently small. In other words, segregation will be stable as income distributions converge in metropolitan areas with significant black populations, but need not be stable in areas with small black populations. This is broadly consistent with empirical realities. Speaking of the decline in segregation during the 1980s, Farley and Frey (1994, p. 40) observe that "the largest declines occurred in metropolitan areas in which blacks made up a small percentage of the neighborhood of the typical white." While this finding has commonly been attributed to the hypothesis that whites are threatened by large numbers of black households in their neighborhoods, our analysis suggest an alternative interpretation. When the share of black households in a city is small, whites sort themselves more extensively by income. The difference in income between more affluent white neighborhoods and black neighborhoods is therefore greater, tempting the highest income black households to move to overwhelmingly white neighborhoods. Thus segregation is less likely to remain stable in cities with small black populations as racial income disparities decline.

Stability of complete segregation does not imply that integrated equilibria cannot also be stable, as the following example illustrates:

Example 3 *Suppose that the income distributions are uniform with support* $[0, 0.9]$ *for black households and* $[0.1, 1]$ *for white households, with all other specifications as in Example 1. Then there exist four equilibria:*

z^*	$\tilde{y}_b$	β_1	β_2	$\bar{y}_1$	$\bar{y}_2$	ρ
0.18	0.90	0.90	0.00	0.42	0.59	0.04
0.24	0.83	0.83	0.07	0.37	0.64	0.06
0.49	0.53	0.53	0.37	0.28	0.73	0.18
0.89	0.03	0.03	0.87	0.48	0.53	0.01

Of the four equilibria identified in the example, only the first (involving segregation) and the third (involving substantial integration) are stable. This can be seen in Figure 3, where the marginal bid-rent functions are depicted. Both stable equilibria are of the kind identified in Proposition 2, with the marginal black household having higher income than the marginal white household.

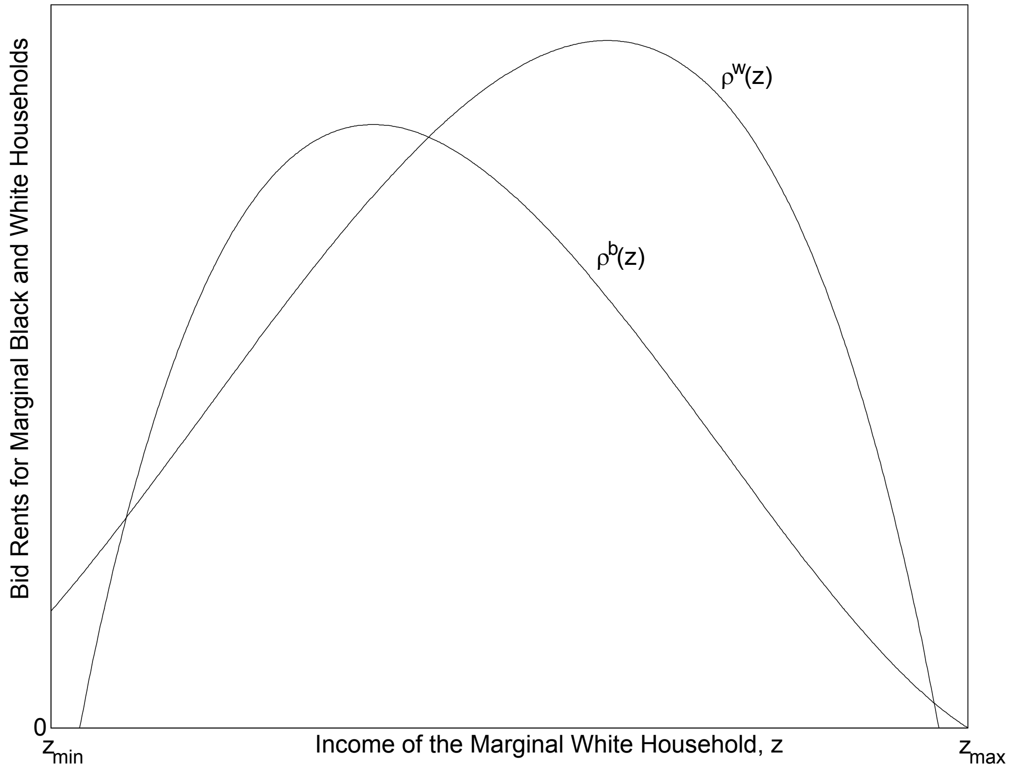

Figure 3: Multiple Stable Equilibria

The above example is robust in that when racial income disparities are sufficiently small, and preferences are sufficiently pro-integrationist, a stable integrated equilibrium always exists:

Proposition 4 *There exist* $\bar{\eta} > 0$ *and* $\bar{\alpha} < 1$ *such that if* $\eta < \bar{\eta}$ *and* $\alpha > \bar{\alpha}$ *both hold, a stable integrated equilibrium exists.*

In combination with Proposition 3, one implication of this result is that multiple stable equilibria exist when preferences over neighborhood racial composition are sufficiently pro-integrationist, racial disparities in the distribution of income are small, and the black share of the metropolitan population is sufficiently large. When $\alpha = 1$ the integrated equilibrium involves identical neighborhood racial compositions and complete stratification by income. In the limiting case when such disparities disappear, there is a stable equilibrium in which there is effectively no racial segregation in residential patterns.

Propositions 3 and 4 may be illustrated by looking at a special case. Suppose that the black and white income distributions are both uniform with supports $[0, \frac{1}{2}(1+\alpha)]$ and $[\frac{1}{2}(1-\alpha), 1]$ respectively, with all other specifications as in Example 2. The manner in which the set of stable equilibria varies with racial income disparities α is shown in Figure 4.

When racial income disparities are extreme (α close to zero) complete segregation is the only stable outcome. As racial income disparities narrow there comes a point when the segregated equilibrium loses stability and the unique stable equilibrium involves some degree of mixing. Beyond this point, convergence of incomes goes hand in hand with greater integration. Eventually α crosses a threshold and multiple equilibria arise, with complete segregation becoming stable. Further convergence of incomes can lead to persistent segregation or to increasing integration: depending on which of the equilibria is selected. When the two income distributions are identical ($\alpha = 1$) the two stable equilibria are at polar extremes: one segregated and the other perfectly integrated with the neighborhoods having identical racial compositions.

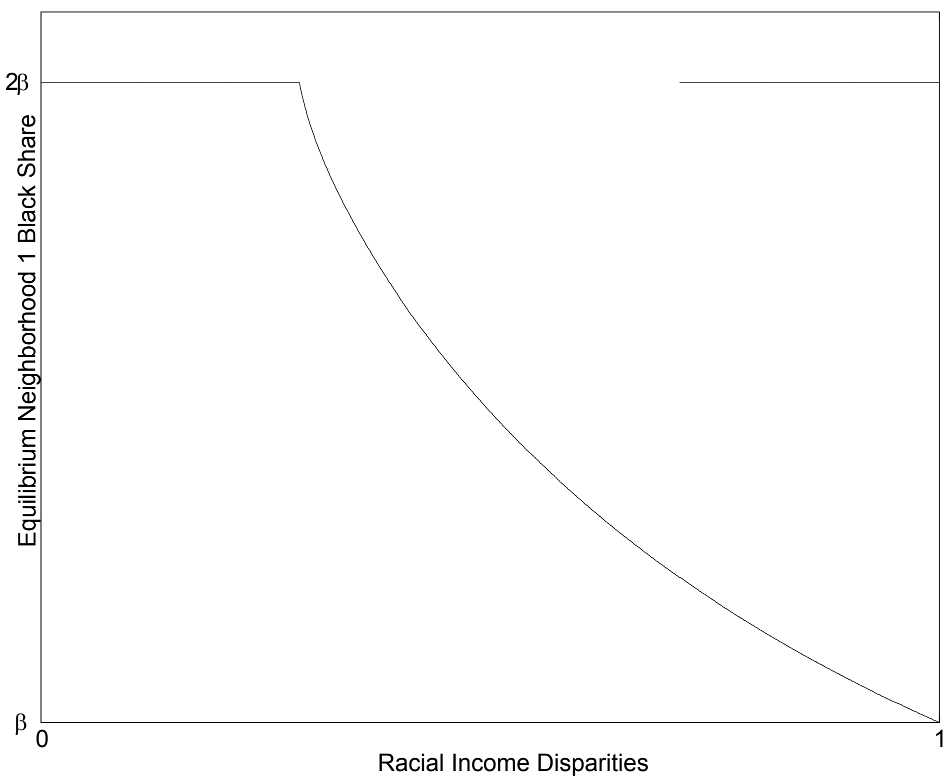

Figure 4: Racial Income Disparities and Equilibrium Segregation

Another implication of Proposition 4 is that stable equilibria can exist which involve higher levels of integration than any household, black or white, considers ideal. For example, if $\eta = \beta = 1/2$, the ideal neighborhood for each household requires that it be in a 75% majority. Yet if $\alpha = 1$, a stable equilibrium exists in which there is complete stratification by income, and exactly half the households in each neighborhood are black. To see why this is stable, suppose that random perturbations cause z to rise so that the second neighborhood is now majority black. Based on preferences over neighborhood composition alone, the marginal black household would outbid the marginal white household for housing in the second neighborhood, leading to cumulative divergence from the equilibrium. But with z exceeding $\frac{1}{2}$, the marginal black household is less affluent than the marginal white household. Based on preferences for neighborhood mean income alone, the latter would outbid the former for housing in the second neighborhood. The combined effect of these two forces determines whether integration is stable. If preferences over neighborhood racial composition are not too strong, the latter effect dominates and integration is stable.

The results in this section suggest that racial disparities in the distribution of income play a subtle and important role in determining patterns of segregation. Even when preferences are strongly pro-integrationist and the ideal neighborhood for all individuals is close to perfectly mixed, complete segregation can result if racial income disparities are negligible or extreme. Multiple equilibria are inevitable when racial income disparities are small. The existence of multiple equilibria suggests that although stable integration may become viable as racial income disparities lessen, a city may remain trapped in the basin of attraction of the segregated equilibrium due to historical patterns of segregation. This is where social policy may be most effective: temporary incentives for segregation may give rise to permanent effects.

In the simple model considered here, the integrated equilibrium results in greater social surplus if preferences are sufficiently pro-integrationist. Integration comes at the cost of greater stratification, however, and this conclusion need not hold once the efficiency effects of stratification are accounted for. These effects can arise because the quality of one's neighborhood affects incentives for human capital accumulation, and hence also aggregate production and growth (Benabou 1996). There are conditions under which stratification by income can be efficiency-reducing, in which case integrationist policies need not be unambiguously welfare improving even when preferences are pro-integrationist. In addition, a shift to an equilibrium with greater integration (and correspondingly greater stratification) lowers neighborhood quality in the poorest neighborhood, which consists disproportionately of black households. The movement of upper-income black households to more affluent communities worsens the conditions for those left behind; a point that has been emphasized by Wilson (1987).[8]

[8]While the effects of greater integration are theoretically ambiguous, Cutler and Glaeser (1997) present evidence supporting the claim that black households overall benefit from declines in segregation.

Finally, our results imply that one cannot expect a narrowing of racial income disparities to lead inevitably to lower segregation. While the convergence of incomes might imply greater integration at integrated equilibria, it may also cause segregated allocations to become stable. From a cross-sectional perspective, cities with lower levels of racial inequality need not be the least segregated. And from a historical perspective, the march towards greater integration may be halted and reversed in some cities as racial inequality *declines*.

7 Conclusions

Given the nation's long history of slavery and *de jure* segregation, race has a degree of salience in American life that perhaps exceeds that of any other socially designated attribute. Although racial attitudes will doubtless continue to evolve, preferences over neighborhood racial composition will remain an important factor affecting household location decisions in the foreseeable future. Such preferences are reflected in the patterns of segregation and stratification that prevail, but in subtle and sometimes unexpected ways.

The main results of our analysis may be summarized as follows. When the black share of the metropolitan population is sufficiently large and racial income disparities are not too extreme, there exist stable equilibria in which black households experience lower neighborhood quality conditional on income than white households. This does not arise from overt discrimination, and is consistent with a desire on the part of all households to live in integrated communities. Segregated allocations are stable when racial income disparities are either very great or very small, but may be unstable when racial inequality lies in some intermediate range. And multiple equilibria arise when racial income disparities are narrow, with extreme segregation and high levels of integration both being consistent with stability.

Our analysis is abstract enough to permit other interpretations. If, instead of race, the dominant attribute governing location decisions were linguistic preference, religious affiliation, or any other observable trait, our basic findings would continue to apply. It is also not necessary to interpret neighborhoods in a spacial sense: interaction in clubs or other voluntary associations will be subject to the same kind of dynamics. One could consider wealth rather than income disparities, and owned rather than rented housing without substantive modification to the model.

The most obvious significant extension of this work would be to allow for the endogenization of income distributions in an intergenerational context. The consideration of a broader class of preferences, intrinsic differences in neighborhood quality, and multiple racial or ethnic groups is also potentially important. Exploring these extensions will yield answers to new questions, generate additional insights, and determine the extent to which the basic insights emerging in this paper hold under more general conditions.

Appendix

Proof of Proposition 1

Let (z^*, ρ^*) be an equilibrium and suppose first that $\rho^* = 0$. Then if any household strictly prefers a neighborhood so do all households belonging to the same race. At any integrated equilibrium there must be black and white households in each neighborhood, so all households must be indifferent between the two neighborhoods. This implies

$$\bar{y}_1(z^*) + v(\beta_1(z^*)) = \bar{y}_2(z^*) + v(\beta_2(z^*)), \tag{2}$$
$$\bar{y}_1(z^*) + v(\omega_1(z^*)) = \bar{y}_2(z^*) + v(\omega_2(z^*)) \tag{3}$$

Using (1), and the fact that $\omega_j = 1 - \beta_j$ we obtain for any allocation z

$$[v(\beta_1(z)) - v(\beta_2(z))] - [v(\omega_1(z)) - v(\omega_2(z))] = 2\eta(\beta_1(z) - \beta_2(z)). \tag{4}$$

In particular, this is true for $z = z^*$ so (2–3) both hold only if $\beta_1(z^*) = \beta_2(z^*)$ and $\bar{y}_1(z^*) = \bar{y}_2(z^*)$. Hence if $\rho^* = 0$ equilibrium allocation must be either segregated or symmetric.

Next suppose $\rho^* \neq 0$ and, without loss of generality, let $\rho^* > 0$. Then for any household with income y in neighborhood 2,

$$u(y) + \bar{y}_1(z^*) + v(r_1(z^*)) \leq u(y - \rho^*) + \bar{y}_2(z^*) + v(r_2(z^*)),$$

where $r = \beta$ if the household is black and $r = \omega$ otherwise. Since u is strictly concave, all households of the same race but with income greater than y strictly prefer neighborhood 2. Hence if $\rho^* \neq 0$ the equilibrium is intraracially stratified. ■

Proof of Proposition 2

Since $\tilde{y}^b(z_{\min}) = y^b_{\max} > z_{\min}$ and $\tilde{y}^b(z_{\max}) = y^b_{\min} < z_{\max}$, the continuity and strict decreasingness of $\tilde{y}^b(z)$ imply a unique $\hat{z}$ which satisfies $\tilde{y}^b(\hat{z}) = \hat{z}$. Since $F^b(\hat{z}) > F^w(\hat{z})$, we must have $\beta_1(\hat{z}) > \beta_2(\hat{z})$. The marginal bid-rents at $\hat{z}$ are defined by the indifference conditions

$$u(\hat{z}) + \bar{y}_1(\hat{z}) + v(\omega_1(\hat{z})) = u(\hat{z} - \rho^w(\hat{z})) + \bar{y}_2(\hat{z}) + v(\omega_2(\hat{z}))$$
$$u(\hat{z}) + \bar{y}_1(\hat{z}) + v(\beta_1(\hat{z})) = u\left(\hat{z} - \rho^b(\hat{z})\right) + \bar{y}_2(\hat{z}) + v(\beta_2(\hat{z}))$$

Since $\beta_1(\hat{z}) > \beta_2(\hat{z})$, (4) implies that $v(\beta_1(\hat{z})) - v(\beta_2(\hat{z})) > v(\omega_1(\hat{z})) - v(\omega_2(\hat{z}))$. This in turn implies that the above indifference conditions can only hold if $\rho^b(\hat{z}) < \rho^w(\hat{z})$.

Suppose first that $\rho^b(z) < \rho^w(z)$ for all $z \in (z_{\min}, \hat{z}]$. Then there is a stable equilibrium at $(z_{\min}, \rho^w(z_{\min}))$ if and only if $\rho^w(z_{\min}) \geq 0$. To see that $\rho^w(z_{\min}) \geq 0$ holds, note that $\omega_1(z_{\min}) = 1 - 2\beta$ and $\omega_2(z_{\min}) = 1$, so from the definition of ρ^w we have

$$u(z_{\min}) + \bar{y}_1(z_{\min}) + v(1-2\beta) = u\left(z_{\min} - \rho^w(z_{\min})\right) + \bar{y}_2(z_{\min}) + v(1). \tag{5}$$

If $\beta > (1-\eta)/2$, then from (1), we have

$$v(1) - v(1-2\beta) = \eta - (1-2\beta)(2\beta+\eta) = 2\beta(2\beta - 1 + \eta) > 0. \tag{6}$$

This, together with (5) and the fact that $\bar{y}_2(z_{\min}) > \bar{y}_1(z_{\min})$, this implies $u\left(z_{\min} - \rho^w(z_{\min})\right) < u(z_{\min})$ and therefore $\rho^w(z_{\min}) > 0$. Hence $(z_{\min}, \rho^w(z_{\min}))$ is a stable equilibrium. Since $\tilde{y}^b(z_{\min}) = y^b_{\max} > z_{\min}$ the result follows.

Finally, suppose that $\rho^b(z) \geq \rho^w(z)$ for some $z \in (z_{\min}, \hat{z})$. Then, since $\rho^b(\hat{z}) < \rho^w(\hat{z})$, the continuity of $\rho^b(z)$ and $\rho^w(z)$ implies that there exists some $z^* \in (z_{\min}, \hat{z})$ such that $\rho^b(z^*) = \rho^w(z^*)$. Let $\rho^* = \rho^b(z^*) = \rho^w(z^*)$. The pair (z^*, ρ^*) is an equilibrium if and only if $\rho^* \geq 0$. Suppose, by way of contradiction, that $\rho^* < 0$. Then the indifference conditions for marginal households may be written

$$\begin{aligned} u(z^*) + \bar{y}_1(z^*) + v(\omega_1(z^*)) &= u(z^* + |\rho^*|) + \bar{y}_2(z^*) + v(\omega_2(z^*)) \\ u(\tilde{y}^b(z^*)) + \bar{y}_1(z^*) + v(\beta_1(z^*)) &= u\left(\tilde{y}^b(z^*) + |\rho^*|\right) + \bar{y}_2(z^*) + v(\beta_2(z^*)) \end{aligned}$$

Since $\beta_1(z^*) > \beta_2(z^*)$ we have $v(\beta_1(z^*)) - v(\beta_2(z^*)) > v(\omega_1(z^*)) - v(\omega_2(z^*))$ from (4). Hence

$$u(z^* + |\rho^*|) - u(z^*) < u\left(\tilde{y}^b(z^*) + |\rho^*|\right) - u(\tilde{y}^b(z^*)),$$

which, since u is concave, implies that $z^* > \tilde{y}^b(z^*)$. But since $z^* < \hat{z}$ and $\tilde{y}^b(z)$ is strictly decreasing we must have $\tilde{y}^b(z^*) > z^*$, a contradiction. Therefore $\rho^* \geq 0$ and (z^*, ρ^*) is an equilibrium. ■

Proof of Proposition 3

First consider part (a). Define $\Delta(y, \rho) \equiv u(y) - u(y - \rho)$. Since u is strictly concave, Δ is strictly decreasing in y whenever $\rho > 0$, and strictly increasing in ρ. The conditions defining the bid-rents for the marginal households are therefore

$$\Delta(z, \rho^w) + \bar{y}_1(z) - \bar{y}_2(z) + v(\omega_1) - v(\omega_2) = 0 \tag{7}$$

$$\Delta(\tilde{y}^b(z), \rho^b) + \bar{y}_1(z) - \bar{y}_2(z) + v(\beta_1) - v(\beta_2) = 0 \tag{8}$$

Consider the pair $(z^*, \rho^*) = (z_{\min}, \rho^w(z_{\min}))$. Since z^* is segregated, $\beta_1(z^*) = 2\beta$, $\beta_2(z^*) = 0$, and $\tilde{y}^b(z^*) = y^b_{\max}$. Subtracting (8) from (7) and using (4) therefore yields

$$\Delta(z^*, \rho^*) - \Delta(y^b_{\max}, \rho^b(z^*)) > 4\beta\eta \tag{9}$$

for all α. Since $\beta > (1-\eta)/2$,we have $v(1) - v(1-2\beta) > 0$ from (6). This, together with the fact that $\bar{y}_2(z_{\min}) > \bar{y}_1(z_{\min})$, implies that $\rho^* = \rho^w(z_{\min}) > 0$. By assumption, $\lim_{\alpha\to 0} y^b_{\max}(\alpha) = \lim_{\alpha\to 0} y^w_{\min}(\alpha) \le z_{\min}$. Since $\rho^* > 0$, this implies

$$\lim_{\alpha\to 0} \left(\Delta(z^*, \rho^*) - \Delta(y^b_{\max}, \rho^*)\right) \le 0$$

Suppose, by way of contradiction, that $\lim_{\alpha\to 0} \rho^b(z^*) \ge \rho^*$. Then

$$\lim_{\alpha\to 0} \Delta(y^b_{\max}, \rho^*) \le \lim_{\alpha\to 0} \Delta(y^b_{\max}, \rho^b(z^*))$$

so that

$$\lim_{\alpha\to 0} \left(\Delta(z^*, \rho^*) - \Delta(y^b_{\max}, \rho^b(z^*))\right) \le 0$$

But this contradicts (9). Hence $\lim_{\alpha\to 0} \rho^b(z) < \rho^*$ and there must exist $\alpha_h > 0$ such that $\rho^* > \rho^b(z^*)$ for all $\alpha < \alpha_h$. For all $\alpha < \alpha_h$, therefore, (z^*, ρ^*) is a stable segregated equilibrium.

Next consider part (b). When $\beta = \frac{1}{2}$ and $\alpha = 1$, $\bar{y}_1(z_{\min}) = \bar{y}_2(z_{\min})$. Since mean neighborhood incomes are continuous in β and α, for any $\varepsilon > 0$ there exist $\beta'_h < \frac{1}{2}$ and $\alpha_h < 1$ such that $\bar{y}_2(z_{\min}) - \bar{y}_1(z_{\min}) < \varepsilon$ for all $\alpha > \alpha_h$ and $\beta > \beta'_h$. From (1) we obtain

$$\begin{aligned} v(1) - v(1-2\beta) &= 2\beta\eta - 2\beta(1-2\beta) \\ v(0) - v(2\beta) &= -2\beta\eta - 2\beta(1-2\beta) \end{aligned}$$

Hence there exists $\beta''_h < \frac{1}{2}$ such that for all $\beta > \beta''_h$, $v(1) - v(1-2\beta) > \beta\eta$ and $v(0) - v(2\beta) < -\beta\eta$. Set $\varepsilon = \beta\eta$ and $\beta_h = \min\{\beta'_h, \beta''_h\}$. Then for all $\alpha > \alpha_h$ and $\beta > \beta_h$ we have $\Delta(z^*, \rho^*) > 0 > \Delta(y^b_{\max}, \rho^b(z^*))$. This implies $\rho^* > 0 > \rho^b(z^*)$, so (z^*, ρ^*) is a stable, segregated equilibrium. ■

Proof of Proposition 4

Suppose $\alpha = 1$ and consider the allocation z with $\tilde{y}^b(z) = z$ (complete stratification). The indifference conditions are

$$u(z) - u(z - \rho^w(z)) + \bar{y}_1(z) - \bar{y}_2(z) = v(1-\beta_2) - v(1-\beta_1) \tag{10}$$

$$u(\tilde{y}^b(z)) - u(\tilde{y}^b(z) - \rho^b(z)) + \bar{y}_1(z) - \bar{y}_2(z) = v(\beta_2) - v(\beta_1) \tag{11}$$

Since complete stratification maximizes mean income differentials, $\bar{y}_1'(z) - \bar{y}_2'(z) = 0$. Hence

$$\begin{aligned} u'(z) - u'(z - \rho^w)\left(1 - \frac{d\rho^w}{dz}\right) &= v'(1-\beta_2)\beta_1' - v'(1-\beta_2)\beta_2' \\ u'\left(\tilde{y}^b(z)\right)\frac{d\tilde{y}^b}{dz} - u'(\tilde{y}^b(z) - \rho^b)\left(\frac{d\tilde{y}^b}{dz} - \frac{d\rho^b}{dz}\right) &= v'(\beta_2)\beta_2' - v'(\beta_1)\beta_1' \end{aligned}$$

Since the income distributions are identical at $\alpha = 1$ we have $\beta_1 = \beta_2 = \beta$. This implies from (10-11) that $\rho^w(z) = \rho^b(z) > 0$ so z is an equilibrium. Let ρ denote the equilibrium rent at z. Then, since $v'(r) = 1 - 2r + \eta$ and $\tilde{y}^b(z) = z$, we have

$$\begin{aligned} u'(z) - u'(z - \rho)\left(1 - \frac{d\rho^w}{dz}\right) &= (2\beta - 1 + \eta)(\beta_1' - \beta_2'), \\ u'(z)\frac{d\tilde{y}^b}{dz} - u'(z - \rho)\left(\frac{d\tilde{y}^b}{dz} - \frac{d\rho^b}{dz}\right) &= (1 - 2\beta + \eta)(\beta_2' - \beta_1'), \end{aligned}$$

Hence

$$u'(z - \rho)\left(\frac{d\rho^b}{dz} - \frac{d\rho^w}{dz}\right) = \left(u'(z - \rho) - u'(z)\right)\left(\frac{d\tilde{y}^b}{dz} - 1\right) + 2\eta(\beta_2' - \beta_1') \qquad (12)$$

The equilibrium rent ρ is determined implicitly by

$$u(z) - u(z - \rho) + \bar{y}_1(z) - \bar{y}_2(z) = 0, \qquad (13)$$

and so ρ is independent of η. This, together with (12) and the fact that $\tilde{y}^b$ is strictly decreasing in z, implies that there exists $\bar{\eta} > 0$ such that for all $\eta < \bar{\eta}$,

$$\frac{d\rho^b}{dz} - \frac{d\rho^w}{dz} < 0. \qquad (14)$$

Hence for all $\eta < \bar{\eta}$ the equilibrium is stable. To complete the proof we need only show that if an integrated equilibrium is stable for $\alpha = 1$, there must exist $\bar{\alpha} < 1$ such that a stable and integrated equilibrium exists for all $\alpha > \bar{\alpha}$. But this follows from the continuity of the marginal-bid rent functions in α. ■

References

[1] Benabou, R. (1992). "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics* 108: 619-52.

[2] Benabou, R. (1996). "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance." *American Economic Review* 86: 584-609.

[3] Bobo, L., H. Schuman, and C. Steeh (1986). "Changing Attitudes Toward Residential Segregation." In *Housing Desegregation and Federal Policy*, edited by J.M. Goering. Chapel Hill: University of North Carolina Press.

[4] Bond, E.W and E.N. Coulson (1889). "Externalities, Filtering, and Neighborhood Change" *Journal of Urban Economics* 26: 231-49.

[5] Council of Economic Advisors (1998). *The Economic Report of the President*. Washington: U.S. Government Printing Office.

[6] Cutler, D.M. and E.L. Glaeser (1997). "Are Ghettos Good or Bad?" *Quarterly Journal of Economics* 112: 827–872.

[7] Cutler, D.M., E.L. Glaeser and J.L. Vigdor (1999). "The Rise and Decline of the American Ghetto." *Journal of Political Economy* 107: 455–506.

[8] de Bartolome, C.A.M. (1990). "Equilibrium and Inefficiency in a Community Model with Peer Group Effects." *Journal of Political Economy* 98: 110-33

[9] Denton, N.A. and D.S. Massey (1988). "Residential Segregation of Blacks, Hispanics and Asians by Socioeconomic Status and Generation." *Social Science Quarterly* 69: 797–817.

[10] Durlauf, S.N. (1996). "A Theory of Persistent Income Inequality". *Journal of Economic Growth* 1: 75-93.

[11] Ellen, I.G. (2000). *Sharing America's Neighborhoods: The Prospects for Stable Racial Integration*. Cambridge: Harvard University Press.

[12] Epple, D. and G.J. Platt (1998). "Equilibrium and Local Redistribution in an Urban Economy When Households Differ in Both Preferences and Incomes." *Journal of Urban Economics* 43: 23-51.

[13] Epple, D. and T. Romer (1991). “Mobility and Redistribution” Journal of Political Economy 99: 828-58.

[14] Farley, R., M. Krysan, T. Jackson, C. Steeh, and K. Reeves (1993). “Causes of Continued Racial Residential Segregation in Detroit: ‘Chocolate City, Vanilla Suburbs’ Revisited” *Journal of Housing Research 4*: 1-38.

[15] Farley, R., H. Schuman, S. Bianci, D. Colasanto, and S. Hatchett (1978). “‘Chocolate City, Vanilla Suburbs’: Will the Trend Toward Racially Separate Communities Continue?” *Social Science Research* 7: 319-44.

[16] Farley, R., C. Steeh, M. Krysan, T. Jackson, and K. Reeves (1994). “Stereotypes and Segregation: Neighborhoods in the Detroit Area.” *American Journal of Sociology* 100: 750-80.

[17] Farley, R. and W.H. Frey (1994). “Changes in the Segregation of Whites from Blacks during the 1980: Small Steps toward a more Integrated Society.” *American Sociological Review* 59: 23–45.

[18] Fernandez, R. and R. Rogerson (1996). “Income Distribution, Communities, and the Quality of Public Education.” *Quarterly Journal of Economics* 111: 135-64.

[19] Glaeser, E.L. and J.L. Vigdor (2001). “Racial Segregation in the 2000 Census: Promising News.” *The Brookings Institution Survey Series,* April 2001. The Brookings Institution: Washington D.C.

[20] Kern, C.R. (1981). “Racial Prejudice and Residential Segregation: The Yinger Model Revisited” *Journal of Urban Economics* 10: 164-72.

[21] Lewis Mumford Center for Comparative Urban and Regional Research (2001). *Metropolitan Racial and Ethnic Change: Census 2000 Data.* State University of New York at Albany (http://www.albany.edu/mumford/census).

[22] Loury, G.C. (1977). “A Dynamic Theory of Racial Income Differences.” In *Women, Minorities and Employment Discrimination*, edited by Phyllis Wallace and Annette LaMond.

[23] Lundberg, S. and R. Startz (1998). “Race, Information, and Segregation.” Mimeo, University of Washington.

[24] Massey, D.S. and N.A. Denton (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge: Harvard University Press.

[25] O'Flaherty, B. (1999). "Troubled Transactions and their Consequences: Race in the United States." Mimeo, Columbia University.

[26] Schelling, T.C. (1971). "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1: 143–86.

[27] Schelling, T.C. (1972). "A Process of Residential Segregation: Neighborhood Tipping." In *Racial Discrimination in Economic Life*, edited by A.H. Pascal. Lexington, MA: Lexington Books.

[28] Schelling, T.C. (1978). *Micromotives and Macrobehavior*. New York: W.W. Norton.

[29] Schnare and McRae (1978). "The Dynamics of Neighborhood Change." *Urban Studies* 15: 327-331.

[30] Schuman, H., C. Steeh, L. Bobo, and M. Krysan (1997). *Racial Attitudes in America: Trends and Interpretations*, Revised Edition. Cambridge: Harvard University Press.

[31] U.S. Census Bureau (2000). *Current Population Reports, P60-209, Money Income in the United States: 1999.* Washington: U.S. Government Printing Office.

[32] Wilson, W.J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

[33] Yinger, J. (1976). "Racial Prejudice and Racial Residential Segregation in an Urban Model." *Journal of Urban Economics* 1: 21-47.

[34] Yinger, J. (1995). *Closed Doors, Opportunities Lost: The Continuing Costs of Housing Discrimination*. New York: Russell Sage Foundation.

[35] Zubrinsky Charles, C. (2001). "Processes of Racial Residential Segregation." In *Urban Inequality: Evidence from Four Cities*, edited by Alice O'Connor, Chris Tilly, and Lawrence D. Bobo. New York: Russell Sage Foundation.