WORKING PAPER

2004

Department of Economics
Tufts University
Medford, MA 02155
(617) 627 – 3560
http://ase.tufts.edu/econ

hudpfinl.tex: February 28, 1999

INCOME MIXING AND HOUSING IN U.S. CITIES: EVIDENCE FROM NEIGHBORHOOD CLUSTERS OF THE AMERICAN HOUSING SURVEY [1]

Anna Hardman and Yannis M. Ioannides

Department of Economics

Tufts University

Medford, MA 02155

Abstract

Most American households live in neighborhoods where the degree of income mixing (and segregation by income) is determined by the housing market without public intervention. The paucity of locational information in micro data sets which have samples representative of the U.S. population has limited our understanding of the extent of market-driven mixing or segregation by income in neighborhoods smaller than census tracts. This paper describes the extent of sorting and mixing by income in the immediate neighborhoods surrounding U.S. urban households in 1985 and 1993, using a unique data set describing the characteristics of the closest neighbors of a representative sample of U.S. urban households.

[1]This project was sponsored by HUD Office of Policy Development and Research Grant No. HP962786. We are grateful to Dr. Kathryn P. Nelson, who initiated this project, for her patient guidance and support, and for her thorough comments and suggestions at each stage of the work. We thank Linda Datcher Loury for her useful comments and Anne Kinsella for excellent research assistance. Additional support by the National Science Foundation, the MacArthur Foundation and Tufts University to Yannis M. Ioannides is gratefully acknowledged.

1 Introduction

Mixed-income housing has been a recurring theme of Federal housing policies and strategies. The 1937 Housing Act, Section 8, created a Federal program " ... for the purpose of promoting economically mixed housing." More recently, the landmark *Mount Laurel* ruling [2] established that a municipality "cannot foreclose persons of low and moderate income from housing and must affirmatively afford that opportunity."

The interest in mixed-income housing has been grounded in the belief that aiding the dispersion of poor households and undoing the heavy concentration of low-income households in dense urban housing developments is socially beneficial. The belief rests on two premises: first, that greater mixing will directly improve living conditions for the poor; second, that it can help break a vicious circle of persistent poverty that feeds antisocial behavior and is more likely to arise where many poor households are concentrated together (and isolated from the rest of the population).

The policy community has increasingly emphasized flexible policy interventions, programs that subsidize households rather than structures. Both HUD and Congress have taken the view that policies that result in heavy concentrations of very low income households in densely built urban areas such as urban public (and particularly project-based) housing are to be avoided. Unlike the urban renewal interventions of the 1950's and 1960's, current policies set out to use federal resources in ways that encourage mixing of households of different incomes as well as races, in the hope that the development of community ties among heterogeneous households will help break down social and ethnic barriers and enrich the urban social environment.[3] It is hoped that such policies will bring about a reversal of "epidemic" forces of urban decay and hence set in motion a process of neighborhood economic and social recovery [4].

Mixed-income housing thus seems to offer direct and indirect benefits (improving the living conditions of the urban poor and neighborhood economic development), which are obviously desirable outcomes of policy intervention. As yet little research has been available to inform housing

[2] *Southern Burlington County NAACP v. Township of Mount Laurel,* 336 A 2nd. 713 (NJ 1975) and 456 A 2nd. 390 (NJ 1983), quoted in de Bartolomé and Rosenthal (1996a).

[3] See Bratt (1989), p. 336-338, for a discussion of how mixed-income housing may help remove "the stigma associated with living in a 'project'."

[4] The so-called "epidemic" theories of ghettos [Jencks and Mayer (1990); Crane (1991)] have recently been cited by the popular press as explanations for the reduction in urban crime in certain cities; see Gladwell (1996).

policy-makers as to the extent of income diversity or mixing that arises without deliberate public intervention. That information would make it much easier to design and evaluate the outcomes of government interventions aimed at improving income diversity in subsidized housing.

A baseline of information has been needed to guide those policies, in the form of detailed description of the extent of income mixing in U.S. neighborhoods. This paper does just that. It summarizes new evidence of the extent of income mixing which develops as a market outcome in U.S. urban areas, at the scale of the immediate neighborhood of urban households. To do so it uses a unique data set, the American Housing Survey's neighborhood clusters data.

The remainder of the paper is organized as follows. Section 2 discusses conceptual issues that are deemed as crucial in understanding socioeconomic characteristics in the pattern of households' location in space. Section 3 describes the American Housing Survey clusters data set in more detail and reviews the availability of variables, especially geographic detail and income variables, which are used later on in the description of income mixing. Section 4 introduces alternative measures of mixing. Section 5 presents our findings, starting with the evidence on income mixing, and then going on to characterize neighborhoods conditional on specific characteristics of kernels, including comparisons between neighborhoods of subsidized and unsubsidized renters, and between families with and without children. Section 6 briefly sets our findings in the context of the existing literature . Section 6 concludes.

2 Conceptual Issues in Spatial Socioeconomic Outcomes

Average housing standards – and more generally, housing standards for most of the population – have increased substantially in the past four decades in the United States, and that increase is well documented. Nevertheless, the housing conditions of low income households, and particularly of low income renters and of the poorest households in the population, have lagged, and remain a major concern of policy makers at all levels of government.

Segregation and mixing by race is a well known and widely discussed feature of the urban landscape in the United States. Patterns of racial segregation and mixing have long been a concern of scholars and policy makers alike. Income segregation has received much less attention until recently, and is much less well understood, both theoretically and empirically [See Vandell (1995) and Ioannides (1996)]. There is growing concern both that segregation by income in U.S. urban areas is growing and that the negative effects of segregation of the urban poor are both more pervasive and more insidious than was previously suspected [see Galster and Killen (1985), Case and Katz (1991) and Abramson, Tobin and VanderGoot (1985)][5].

The main source of both rental and owner-occupied housing services in the United States is the housing market. The distribution of housing outcomes reflects the distribution of income, but several qualifications are in order. At high income levels housing consumption increases with income, but less than proportionately to income. As a result, housing consumption is less unequally distributed than income. Housing markets are highly imperfect, moreover, and it would be overly facile to attribute all the remaining misallocation and inequity in housing consumption to fundamental inequities in the income distribution. Some at least of the misallocation and inequities we observe in the housing sector can be attributed to market failures which are appropriately addressed by policy interventions. To design feasible, cost-effective policy interventions, we need to understand the operation of the market and to know what kind of mixing or sorting by income it produces.

Residential segregation or mixing (heterogeneity) by income is a spatial outcome of the operation of the housing market. At the same time, residential neighborhoods are widely recognized as a source of one important type of market failure: externalities, both positive and negative [Galster and Killen (1995)]. We know that income differences and similarities, within and between neighborhoods, contribute significantly to neighborhood interactions and residents' outcomes. For low income households in particular, housing market outcomes can magnify income inequality in both static and dynamic ways. Mixing of income levels within neighborhoods has been found to be an important (and positive) influence on the socioeconomic opportunities which are available to neighborhood residents. Neighbors' characteristics affect the choices visible to and made by households, and youths in particular, living in those neighborhoods.

Because the distribution of income within different ethnic groups varies widely across groups, income and race segregation are highly correlated [Massey and Denton (1993)]. Segregation by income and race can magnify income inequality. Some of these effects can be seen as negative externalities of segregation of the poor. The consequences of discrimination and segregation in

[5]Recent research with data from European countries appears to confirm recent patterns of urban settlement there similar to those in the United States, with growing segregation by income [Huttman *et al.* (1991)].

housing markets are exacerbated because housing outcomes directly affect access to jobs. Both institutional factors such as restrictive zoning and discrimination in the market for housing, and market forces [Vandell (1995)] can mean that good jobs are accessible only with high transportation costs if at all (the so-called spatial mismatch of jobs).

The basic economics of spatial outcomes in urban areas may be outlined as follows. On the demand side of the market, sorting or clustering reflects the distribution among households of incomes and endowments of wealth and skills. It also reflects the distribution among households of both preferences for space and other attributes of housing and preferences for amenities of neighborhoods and communities more generally. A further complicating factor is introduced when neighbors' characteristics themselves influence movers' perceptions of neighborhood amenities. On the supply side of the market, spatial differences in the price of land contribute to the formation of neighborhoods of housing unit types. Together these market forces are a powerful source of spatial bias toward intra-neighborhood homogeneity.[6]

The dynamic consequences of spatial outcomes in the housing market are correspondingly important. In many U.S. metropolitan areas, the structure of local government and of housing markets leads to the concentration of poorer households in poor communities that are unable to invest large amounts of resources in the schooling of their residents. Furthermore, different communities may offer very different access to resources that affect the formation of human capital more generally. Each of these dynamic consequences in turn may perpetuate and even intensify initial differences in income and wealth among households. Over time, this may produce a worsening of the distribution of income and of housing outcomes.

In the light of the static and dynamic consequences of spatial segregation and its importance in public policy, it is perhaps surprising that the extent of income mixing within neighborhoods of the metropolitan areas of the United States has until recently received so little attention, and that it is still relatively poorly understood, at a time when our understanding of the working of the U.S. housing market has grown rapidly. The fact that so little is known about the extent of income mixing in U.S. neighborhoods is largely attributable to a shortage of suitable data. Research on mixing and segregation has been constrained by the limitations imposed by confidentiality on the use of survey

[6] Vandell [1995] provides a valuable and more extensive summary of economic models of residential location as they affect spatial heterogeneity in urban areas.

data. Location-specific data can endanger the confidentiality of individual respondents' answers to surveys. As a result, only very limited neighborhood-specific or locational data are available for the large household-level data sets (such as the Public User Micro Sample of the U.S. census and the American Housing Survey), which are widely used by researchers and have improved our understanding of housing markets.

Research on the socioeconomic and ethnic composition of urban areas has of necessity relied on aggregate data collected within geographically defined but much larger spatial units, such as census tracts with an average population of 2500 to 7000 [White (1987)], Metropolitan Statistical Areas, or states. Other studies use micro data (from the American Housing Survey, or the Panel Study of Income Dynamics, for example) which for reasons of confidentiality lack detailed locational information beyond central city or suburban location within a known metropolitan area. Additional understanding has been obtained from studies using local data [see Galster (1987)]. Both census tracts and MSAs are too large to allow us to study many of the interactions which are most pertinent to income mixing. Moreover, it is informative to be able to contrast findings based on the smallest scale of neighborhood, with those obtained at a higher level of aggregation.

Income mixing and sorting (or clustering and segregation) can of course be viewed at many scales. A relatively neglected aspect of income mixing in U.S. residential neighborhoods is the neighborhood immediately surrounding a household – its closest neighbors. This is a scale which is especially important to the designers and developers of planned mixed-income neighborhoods and to policy makers who wish to influence such decisions.

The importance of this scale is highlighted by the work of Thomas Schelling [Schelling (1971; 1978))] who models the spatial outcomes which are possible where different individuals differ with respect to their tolerance (or preference) for immediate neighbors who are similar to or who differ from them with respect to various characteristics. Schelling's theory describes how the details of how people feel about interaction in their immediate neighborhoods can be responsible for key features of large scale structures in space. Neighbor-to-neighbor interactions can have large-scale consequences, because they can lead to chain reactions. One household moves for its own reasons, but if its move tips some balance, it may cause others to move, too. As a result, a variety of stable and unstable spatial outcomes are possible. Schelling's model provides valuable theoretical underpinnings for the intuitive insight and growing body of evidence that immediate proximity is an important element of the social fabric of U.S. cities. To understand patterns of mixing at larger scales, and to assess the feasibility and potential impact of deliberate policies aimed at income mixing, we clearly need also to examine mixing at this much smaller scale.

3 The American Housing Survey Clusters Data

The present study examines empirical evidence of the extent of income mixing at the scale of the immediate neighborhood. It draws on a recent feature of the American Housing Survey (AHS))[7]. Studies using the AHS have largely defined the basic framework for housing policy and research in the United States.[8]. However, as we noted above, confidentiality severely limited the amount of spatial data which could be released. As a result users could not, for example, compare households or dwellings in the sample in the context of data about their neighborhood (census tract or even town).

To provide more information on the social and physical environment of the units sampled, in 1985, 1989 and 1993, about one-hundredth of the dwelling units in the AHS national core urban sample (680 in 1985 and more in subsequent years) were selected. Approximately ten of their nearest neighbors were also interviewed. Confidentiality is not endangered because the location of each cluster in its metropolitan area is not revealed. Nevertheless, this data set is uniquely valuable because it provides an extraordinary view of clusters of immediate neighbors in U.S. metropolitan areas. It constitutes a particularly suitable frame for assessing in micro detail the extent of income mixing in U.S. urban areas. To date, only a small number of studies [9] have made use of the AHS clusters data.

In this study we use data from the 1985 and 1993 clusters data. Each *cluster* is made up of the randomly chosen member of the national AHS sample of urban dwelling units, the *kernel*, and the ten homes closest to it, its *neighbors* [Hadden and Leger (1990), p. 1-51]. The clusters contain

[7]The AHS is a panel data set designed to represent dwellings in the United States; it is made up of a sample of more than 50,000 dwelling units whose residents are interviewed every two years.

[8]Khadduri and Nelson (1992).

[9]See de Bartolomé and Rosenthal (1995); Gabriel and Rosenthal (1995; 1996); Hoyt and Rosenthal (1995).; Ioannides (1996) is the only paper we are aware of using the 1993 clusters data.

fewer than ten observations for neighbors if some could not be interviewed.[10]

There is no unique scholarly or colloquial definition of a neighborhood [11]. In practice, the availability of data explains the widespread use of census tracts or towns as "statistical neighborhoods" [White(1987)]. The information available from the clusters data about the income distribution of the immediate neighbors of sampled dwellings is particularly valuable in understanding the extent of market-driven income mixing in the United States, because it allows us to study a much finer grain of neighborhood than the census tract. However, it is important to realize the limitations inherent in the design of the sample.

The kernels represent the entire population of American urban households; the clusters (neighbors) represent the households living closest to them. The kernels are *not* a sample of neighborhood centers, and the clusters data do not constitute a random sample of neighborhoods. Nevertheless, the clusters are designed to represents the characteristics of the immediate neighbors (and neighborhood in that sense) surrounding the kernels, and the clusters represent the entire population of urban neighbors (or neighborhood *in that sense*). Because the cluster corresponds to the immediate residential neighborhood of the kernel household, in this paper we use "neighborhood" to refer to the place where the kernels and their neighbors live and "neighbors" to refer to the households sampled in the clusters (not including the kernels). We know little about the location of each cluster[12], and nothing about the distance from the clusters to the center of the urban or metropolitan areas where they are located, or about their proximity to the physical features (roads, parks, shopping districts) which often define neighborhoods.

4 Measures of Income Mixing and Sorting in Neighborhoods

Measures of household income and income distribution always need careful interpretation: the same household income represents a different standard of living for a large household than it does for a small one. When we are comparing households across different urban areas, a further complexity

[10]Details of the sample structure and data availability are included in the Appendix.

[11]The theoretical literature certainly does not resolve the issue of how to define a neighborhood. Ellickson (1979) and Goodman (1989) offer theoretical insights, while Ioannides (1996) employs a Nash-equilibrium notion of neighborhood interactions.

[12]The location of the urban clusters examined here is given by metropolitan area (or state for non-metropolitan urban clusters) as metropolitan central city, metropolitan suburban or non-metropolitan urban.

is introduced: the same household income has very different purchasing power in a housing market where prices are high than in one where they are relatively low. Median household incomes (reflecting earnings in the local job market) are a proxy, albeit imperfect, for relative cost of living in different metropolitan areas. In our tables we therefore group households by their household income relative to the appropriate HUD adjusted area median family income (HAMFI). The HUD-adjusted area median family income is based on the median *family* income for the metropolitan area, adjusted for household size. HAMFI is a policy tool used to determine program eligibility for many HUD programs and housing programs administered by other Federal agencies.[13] Here we use it to normalize incomes by household size and income relative to the urban area where the cluster is located. [14]

4.1 Income Categories

Data on income for households needed careful interpretation: the same household income may represent a very different standard of living for a large household than it does for a poor household. In much of the analysis of income mixing in this report we have used household income standardized for household size as well as relative to the appropriate level of HAMFI. To describe income levels our starting point was the income categories used by HUD to define program eligibility cutoffs.

To give a detailed picture of income mixing in our tabulations and analysis we have specified eight categories of incomes, defining them again in terms of HAMFI. Each of those eight categories corresponds very roughly to a little more than a decile of the household income distribution. The names and definition of those categories are summarized in Figure 1 (which includes HUD's categories for comparison).

FIGURE 1 ABOUT HERE

4.2 How Much Mixing? The Benchmarks of Perfect Mixing and Perfect Sorting

The following section describes the patterns of mixing (or sorting) which we observe in U.S. neighborhoods. Most of the measures of segregation used in the literature are designed to describe

[13]The definition and use of the HUD-adjusted area median family income is described in more detail in HUD Office of Policy Development and Research (1996), HUD (1996) for short.

[14]It also has the advantage that the extent of mixing for households in different income groups can be compared with their eligibility for current programs.

situations where a particular area, say a city, is subdivided into smaller areal units, and the extent of segregation is examined across the constituents of the area.[15]

To clarify the outcomes possible and one potential complicating feature of the distribution of income, consider a hypothetical population which consists of two types of people, high-income and low-income. Suppose that the economic system can allocate them into groups of equal size, neighborhoods, in two alternative ways. In one, "perfect sorting" high-income people are completely segregated from low-income people. The rich live in rich neighborhoods, and the poor in poor neighborhoods. Each person is identical to all her neighbors. The numbers of rich and poor neighborhoods will be in exact proportion to the numbers of high income and low income people in the population. The second way of allocating people to neighborhoods is "perfect mixing" in which the proportion of high- and of low-income people in each neighborhood is the same as in the entire population.

A high degree of local within-neighborhood homogeneity (perfect sorting) is compatible with varying degrees of metropolitan, regional or national heterogeneity, depending on the size of the "sorted" neighborhoods. At the same time, widening income dispersion (however measured) within a neighborhood may reflect not a real increase in income mixing within the neighborhood, but increased inequality of incomes in a metropolitan, regional or national context.

Caution is needed in interpreting the overall welfare significance of measures which show apparent increased income mixing at the neighborhood scale, which may simply mirror widening income inequality regionally and nationally.[16] This is particularly relevant in our context, as an increase in the inequality of family incomes and earnings has been observed and widely discussed in the United States since 1973. Specifically, the ratio of the 90th percentile family income to the 10th percentile for the entire U.S. has increased from about 5.5 in 1973 to above 9 in 1994. Similarly

[15]This is done, for example, by the five dimensions proposed as measures of segregation by Massey and Denton (1988). A measure of segregation or geographical dispersion may pertain to first, *evenness*, that is, how does the percentage of poor within each neighborhood compare to the percentage of poor within the metropolitan area. Second, *exposure* of one group to the other which reflects the degree of potential contact between poor and nonpoor through sharing neighborhoods in common. Exposure may be measured in two complementary ways, either through the extent of interaction or isolation. Third, *clustering* describing the extent to which the poor and nonpoor occupy adjoining areas, with maximum clustering taking the form of poor residential areas forming a large contiguous ghetto. Fourth, *centralization* refers to poor households' location close to a metropolitan area center. And, fifth, *concentration*, a measure of relative density in spatial patterns, which refers to the relative amount of space occupied by the two groups.

[16]More generally, the notion of income mixing in relation to the overall income distribution is akin to the notion of decreasing *polarization* [Esteban and Ray (1994)].

pronounced changes have been measured for six of the nine U.S. regions defined by the U.S. Bureau of the Census [Bradbury (1996)].

Another indicator of the complexity inherent in attempting to measure mixing or segregation is the recent evidence of widening income segregation in urban areas at a time when we observe a slight narrowing of racial segregation and an increase in income inequality. Mayer (1996) compares the central city concentration of families in each income quintile to the central city concentration of the average family using Current Population Survey (CPS) data for MSAs from 1964 to 1994. He shows that the ratio of the central city percentage of families in the bottom quintile to the central city percentage of all families exhibits an upward trend and the respective ratio for the top quintile exhibits an downward trend. Our own calculations [17] with published data from the same original data source, the March CPS, find increasing asymmetry in the distribution of family income which suggests increasing income inequality within the central cities of large metropolitan areas in the United States, those with population exceeding 1,000,000 inhabitants.

Here, however, we seek to draw conclusions about the pattern of segregation along several dimensions from a random sample of households in the United States and a cluster of up to ten of their close neighbors. As statistical descriptions of characteristics of neighborhood clusters we have used readily available statistical concepts. It was sometimes necessary to modify them in order to describe variables of interest for entire clusters.

To measure the extent of mixing or sorting, within and between neighborhoods, we examine sample means and coefficients of variation (CV from now on) [18] for variables of interest, including household income, age of the household head, education of the household head, and numbers of children present. The sample mean and CV are calculated for kernel households.[19] We can then

[17] Ioannides (1996) reports the ratio of mean to median income of families in central cities and compares it to those outside central cities of large metropolitan areas, finding changes which are more pronounced in central cities than outside central cities:

Year	1971	1978	1981	1985	1989	1992
Central Cities	1.149	1.19	1.189	1.242	1.265	1.312
Outside Central Cities	1.121	1.128	1.124	1.147	1.183	1.163
Entire U.S.	1.126	1.139	1.154	1.188	1.213	1.208

[18] The coefficient of variation is the standard deviation divided by the mean.

[19] We computed the coefficient of variation of income for each neighborhood in which five or more households reported income.

compare this measure of dispersion for the entire population of urban households with the mean CV for neighbors (a measure of the average amount of dispersion within neighborhoods) and the CV of the CV for neighbors (a measure of the variability between neighborhoods in the amount of dispersion). These measures are estimated for 1985 and 1993 (giving us indications of possible changes in mixing over time). Using these measures we can compare the amount of dispersion within and between neighborhoods for particular subsets of kernels, such as unsubsidized and subsidized renters; we also compare dispersion by region and type of place.

We examine measures of mixing conditional upon kernels' characteristics, as well as their decisions. We measure the average number of neighbors with a particular characteristic, and then compare the overall mean with means for subsets of kernels. For example, consider the percentage of owners among the neighbors in a cluster. How might it vary with the kernel's own characteristics? Does the percentage of owners among a kernel's neighbors depend on whether the kernel is an owner or a renter? On whether she is a subsidized or unsubsidized renter? On whether she is white or nonwhite? Similarly we can ask about the mean age and level of education of neighbors for kernels of different type, or about the proportion of neighbors who are white or nonwhite. These descriptive statistics are useful in understanding the microstructure of U.S. neighborhoods. However simple, they have never been measured before at such a level of disaggregation. They are motivated by Schelling's theory of residential segregation and will be referred to as *Schelling Statistics*.

To provide a benchmark, we can consider three hypothetical cases. The first is the "perfect sorting" described above. When all households are "perfectly sorted", each kernel is living in a neighborhood with neighbors identical with the kernel and with each other. The CV for each neighborhood would then be zero (since the standard deviation is zero); the mean CV for neighbors would be zero, and the CV of the CV for neighbors would also be zero. In the second case, "perfect mixing" as described above, proportions of household types in each neighborhood precisely replicate the entire population. Then the CV within each neighborhood (and the mean CV for neighborhoods) would be the same as the CV for the population, and the CV of the CV for neighbors would be zero (since all neighborhoods have an identical mix of household types). The third case is a mix of the first two: there are two types of neighborhoods, half of them sorted and half mixed. The CV for half the neighborhoods would be zero and for half it would be the same as for the population. The mean CV for neighborhoods would be less than the population CV (since not all neighborhoods are mixed) and the CV of the CV for neighborhoods would be positive (since neighborhoods differ).

Using this approach, we can examine the effect on mixing of the presence or absence of households with specific characteristics, and restrict attention to the neighborhoods of kernels with particular physical and/or economic characteristics. We can ask, for example, whether the neighborhoods of owner-occupier households are more or less heterogeneous than the neighborhoods of unsubsidized renters. Section 5.2 below uses the CV to measure the dispersion in income distribution within and between neighborhoods and Section 5.3 uses means and CV's to analyze the extent of mixing in the neighborhoods of kernels with specified socioeconomic and dwelling characteristics.

In sections 5.2 and 5.4 below, we use quantiles of the distribution of incomes to describe both kernels and their neighbors. We use this approach to describe the likelihood of income mixing, the *probability* that kernels at each income level will live in different types of neighborhoods. Neighborhoods are characterized in terms of the income distribution of the neighbors who make up the cluster. To describe both kernels and neighbors we again use HAMFI based measures, this time characterizing them in terms of quantiles.

One way to describe a neighborhood is in terms of the median income of its residents (or any other quantile of their incomes). Here we first characterize the overall distribution of income in U.S. neighborhoods. We then look at access to these neighborhoods for kernels of different types. For example, we can ask what is the probability that a low income kernel will be living in a low, middle, or upper income neighborhood. We can ask "What is the probability that the low income kernel will have neighbors who have slightly or substantially higher incomes?" We then use a more limited subset of measures (usually the median for a neighborhood) to look at the probability that households with a wider variety of specified characteristics will live in neighborhoods with neighbors with high or low median incomes.

5 Evidence of Income-Mixing in the Immediate Neighborhood of U.S. Households

Households are not, of course, distributed randomly in an urban area. Households' tastes for housing space, quality and amenities combine with their incomes and assets to define their demands for housing of different types in different locations. Those tastes may include preferences for segregation by income or race. Housing choices are further affected by a large number of spatially differentiated factors including the location and accessibility of jobs, and the availability and quality of public and privately provided amenities such as schools, libraries, churches and recreation sites.

Two theoretical models have profoundly influenced economists' thinking about income mixing in U.S. cities. The Tiebout theory [Tiebout (1956)], which implies that households sort themselves into communities of people with similar tastes and incomes, has been widely adopted to explain the relative homogeneity of U.S. suburban communities. Households choose locations which offer the best available match to their preference and ability to pay (via local taxes) for locally produced public services, particularly public education. A second influential theoretical proposition has been the Alonso-Mills-Muth model, with its key prediction that otherwise identical individuals who differ only in terms of income will occupy successive (concentric) zones in a monocentric city.[20] This will produce income segregation related to distance from the Central Business District of the city: the higher the household's income the further away from the center it will locate. Both the Tiebout and the Alonso-Muth-Mills theories in their pure form are highly stylized models of inherently complex phenomena. Both contribute to explaining the formation of patterns of income mixing (or segregation). More important, neither theory explains the complex residential patterns which we observe within the typical metropolitan area in the United States. In what follows, we use the clusters data to describe that complexity.

5.1 The Tables

Our analysis culminates in a number of tables, which are given at the end of the main body of the paper. The first two tables, Tables 1 and 2, use our data to describe two important dimensions of the income distribution when income distribution is measured relative to HAMFI. Table 1 compares

[20]See Muth (1969) for a basic exposition.

the income distribution across the four main U.S. regions, i.e., Midwest, Northeast, South and West; Table 2 compares across central cities, suburban and urban nonmetropolitan areas. Tables 3 and 4 provide a detailed analysis on how mixing with respect to income and other characteristics changes when we condition on the characteristics of the kernels. Tables 5 and 6 report alternative measures of income mixing for kernels in each category of the income distribution. Table 6, in particular, compares the neighborhoods of subsidized from unsubsidized renters. Table 7 provides detail on income mixing in neighborhoods with multi unit structures.[21]

Since most of our statistical inference is based on a sample,[22] it is appropriate to ask how representative it is. We are prevented from obtaining weighted statistics (weights are available for kernels but not for the neighbors). In particular, our data are consistent with the percentiles of the distribution of household income relative to HAMFI.[23]

5.2 Income Distribution

This section describes the distribution of incomes within neighborhoods. One summary measure of neighborhood income is the median income of the cluster in which the kernel lives. Neighborhood median incomes are more tightly distributed (as we would expect if there is any mixing at all) than household incomes.

Comparing the distributions of household (kernel) incomes and neighborhood medians we see some evidence that extremely low and very low income households may be becoming relatively

[21]Additional and more detailed tables are given in the Appendix. Table A.1 provides details on the availability of AHS clusters data, available by Standard Metropolitan Area (SMA). That table also includes data on the HUD-Adjusted Area Median Family Income (HAMFI) for each SMA. Table A.2 gives some basic descriptive statistics for the clusters dataset and compares them with national data. It summarizes data on households and the neighborhoods in which they reside, as made possible by the clusters data, and on dwellings. The data from the entire national sample reported in the bottom of that table allows us to conclude that even though the subsample of the national sample that we are working with is rather small, the overall differences are rather minor. Table A.3 offers additional comparisons between the AHS and national data. Tables A.5–7 in the Appendix offer additional detail along these dimensions.

[22]One which is roughly one-tenth of the national AHS sample and is made up of one hundredth of the national sample plus their ten nearest neighbors.

[23]The descriptive statistics given in Table A.3 in the Appendix show that several key features are consistent with published U.S. data, based on the national AHS sample. In addition, we compare our own computations, reported in Table A.3, with frequencies obtained from the entire national sample of the AHS [HUD (1996), p. B-3]. In 1993, according to HUD's definitions, 15 percent of households have *extremely low incomes,* (17 percent in our own calculations). Similarly, 28 percent (30 percent) of households have *very low incomes,*; 45 percent (48 percent) have *low incomes,* ; 22 percent (21 percent) of households have *middle incomes,*; and 33 percent (31 percent) of households have *upper incomes,*. Our unweighted computations are very close to those based on the full AHS sample, although they tend to overestimate the lowest income category and underestimate the highest ones.

more isolated, while other neighborhoods become more mixed. In 1985, 29 percent of households had median incomes in the extremely low or very low range (less than 50 percent of HAMFI). The neighborhood median fell in that range in 25 percent of all neighborhoods. The corresponding figure for neighborhoods and households in 1993 were 30 percent of households and 29 percent of neighborhoods. The proportion of neighborhoods with median incomes in the very low income range (at least half the residents had extremely low or very low incomes) is evidence of a growing number of relatively homogeneous low income neighborhoods. At the other extreme of the income distribution, 33 percent of households in 1985 and 31 percent in 1993 had incomes in the upper income range (120% of HAMFI or more). The neighborhood median fell in that range in 20 percent of neighborhoods in 1985 and 16 percent in 1993 (Figure 2). The apparent decline in the proportion of neighborhoods with medians in the upper income range may reflect entry of new households or changes in the income of long-term residents.

FIGURE 2 ABOUT HERE

In seeking to identify changes between 1985 and 1993, we need to keep in mind two forces which will be reflected in both household and neighborhood median incomes: first, population movements within and between regions (faster population growth in the West, for example) and second, the widening of the national income distribution.[24] The ratio of mean to median income has increased from 1.231 to 1.326, for the entire United States (our own computations of 1.279 and 1.327 respectively are clearly very close.) The fact that the ratio of mean to median is almost identical in our data to the U.S. figures implies that the same widening of the income distribution which occurred for the U.S. as a whole is (not surprisingly) reflected in the distribution of income for metropolitan areas.

Movement within regions and metropolitan areas probably accounts for some of the changing pattern of dispersion. Table 2 shows differences in the distribution of household incomes, depending on whether the dwelling unit lies in a central city, in a nonmetropolitan urban area, and in a suburb of a central city.[25] Comparing the three columns of Table 2, we can confirm that in both 1985 and

[24]Table A.3 shows the distribution of income in the sample for the four main Census regions, together with supplementary information from the U.S. Statistical Abstract which highlights the phenomenon of widening in the income distribution between 1985 and 1993. Table A.4 shows the distribution of income in the sample for three types of urban neighborhood: metropolitan central city, metropolitan suburban and nonmetropolitan urban.

[25]In constructing this table we aggregated the basic categories as follows. *Central city* corresponds to METRO=1; *suburb* includes METRO=2 (urbanized suburb), METRO=3, (other urban suburb); *urban nonmetropolitan area*

1993 central city neighborhoods have higher concentrations of low income households and suburbs have higher concentrations of high income households.

The second reason for caution in interpreting the changes in the distribution of neighborhood medians shown in Table A.3 is that comparisons based on HAMFI convey a slightly different picture from comparisons based on income alone, primarily because of the nature of the measure we are using. Because HAMFI is designed to adjusts for household size, subtle changes in demographics and population growth along regional dimensions are bound to affect the regional averages. Moreover, of course, HAMFI is established using median family income, but ours is sample is a sample of households and not families.

In examining changes in the dispersion of household income for 1985 and 1993 (Table 3), caution is needed. The median household income is significantly lower than the U.S. median family income. As a result, the relationship of household income to the area median income(HAMFI), is quite different from the relationship of U.S. mean household income to U.S. median income. U.S. median household income was $23618 in 1985 (76.6 percent of $30845, the average value for HAMFI in that same year). In 1993, U.S. median household income was $31241, (75.5 percent of $41365, the average value of HAMFI in our sample for that year). [26]

Similarly, it follows from the definition of HAMFI and the adjustments embodied in it, that HAMFI does not (and should not be expected to) correspond to the true median for our sample of households. In 1993, 52 percent of households in our entire sample have income above 80% of HAMFI, and only 40 percent have incomes greater than HAMFI. As a result, caution is needed in making comparisons across time of incomes defined relative to HAMFI.

The design of HAMFI as an equivalence scale is particularly useful in comparison across regions, as in Table 1, and by type of urban area, as in Table 2, within the same year. The Midwest and the North East exhibit less concentration in the lower part of the distribution, relative to the entire urban U.S., while the opposite is true for the South and to some extent for the West as well. The same broad picture is preserved in 1993. The greater dispersion of the income distribution is

includes METRO=5, 6, 7 (urbanized area, nonmetropolitan other urban).

[26] Comparison of countercumulative distributions for household income relative to HAMFI also indicate an unmistakable shift, from 1985 to 1993, from the upper to the lower part of the distribution, for entire sample as well as for each of the principal Census regions. The countercumulative distribution functions for household income as a percentage of HAMFI for 1993 lie entirely below those for 1985. This means that the percentage of households whose relative income is greater than any particular value is greater in 1985 than in 1993.

reflected in the greater maximum reported household income, although we note that one measure of dispersion, the CV for kernels, fell slightly between 1985 and 1993. [27]

FIGURE 3 ABOUT HERE

The CV for each neighborhood measures dispersion in the income distribution within the neighborhood. Figure 3 shows the range of values and quartiles of the distribution of neighborhood CV's: they ranged from 7.5 to 192.7 in 1985. The change between 1985 and 1993 in the distribution is interesting: in 1993 there are fewer neighborhoods with very large CV (and a smaller maximum) but the CV for the great majority of neighborhoods increased between 1985 and 1993. This is evidence of an increase in mixing or diversity in a wide range of neighborhoods. That increase is particularly striking when we compare it with the decline in the CV for kernels between 1985 and 1993.

The medians and coefficients of variation are summary measures of income distribution within and between neighborhoods. It is also useful, however, to look at how the income distribution within neighborhoods varies for different types of kernel. To interpret the results from the sample, we consider first the two extremes ("perfect mixing" and "perfect sorting") which we used as benchmarks in the previous section. If all households lived in neighborhoods characterized by "perfect mixing" then the distribution of income within each neighborhood would replicate the distribution in the whole population. Figure 4 shows the distribution of income of neighbors for all kernels in a world of "perfect mixing". In all clusters, the neighborhood minimum income falls in the low income range (as it does in the urban population). In all clusters, the neighborhood median income is in the middle income range (again as in the urban population as a whole) and in all clusters the maximum income is in the upper income range (as in the population.

FIGURE 4 ABOUT HERE

Compare that distribution with the outcome with "perfect sorting". Each neighborhood has residents who are identical to each other. In that case (see Figure 5) using the 1985 population proportions as a starting point, 29 percent of clusters would have low income residents (corresponding to the proportion of the population with low incomes); 38 percent of clusters would have residents with middle incomes and 33 percent of the clusters would have residents with high incomes. Within

[27] The distribution of income for kernels ranges from a minimum of a reported income of $0 to a maximum of $210,000 in 1985 and $300,000 in 1993.

each cluster, the minimum, median and maximum income would all fall in the same range of incomes – a cluster with low income residents in a perfectly sorted world, for example, would have its minimum, median and maximum income in the low income range.

FIGURE 5 ABOUT HERE

We can now use the two hypothetical extremes to compare with the actual distribution of incomes for all clusters in 1985 (Figure 6). It falls between the two extreme benchmarks. Most neighborhoods (84.1 percent of clusters) have at least one resident with a low income. With 'perfect mixing' we would expect that figure to be 100 percent; with 'perfect sorting' it would be 29 percent.[28] Upper income households are not as widely dispersed, but 72.5 percent of all neighborhoods have at least one resident in the upper income group (with "perfect sorting" it would be 33 percent). The more detailed information in Figure 2 is consistent with this finding, but suggests that the atypical neighborhood residents may not be numerous: neighborhood median income is in the middle income range defined in our hypothetical examples (50-120 percent of HAMFI) in 55 percent of neighborhoods, compared with the 38 percent we would expect with perfect sorting and the 100 percent expected with perfect mixing.

FIGURE 6 ABOUT HERE

To investigate how the neighborhoods of the poorest urban households compare with U.S. neighborhoods as a whole, we calculated the distribution of income of neighbors in 1985, restricting our sample to kernels with extremely low incomes (Figure 7) and with very low incomes (Figure 8). Because of sample sizes, those results should be interpreted with caution. Those distributions also suggest that while some extremely low income and very low income households live in neighborhoods where moderately low income and middle income households are present, a small proportion of neighborhoods are occupied exclusively by the poorest households. In 14 percent of neighborhoods of extremely poor kernels, the maximum income of neighbors was in the extremely low income or very low income category. In 59 percent of neighborhoods of extremely poor kernels, neighborhood median income was in the very low income or extremely low income category. Another 36 percent of extremely poor kernels lived in neighborhoods with median incomes in the moderately low or middle income categories (predominantly the former). The pattern for kernels with very low incomes is

[28] In Table 6 this is shown as the percentage of clusters the income range of the neighborhood minimum income: in 84.1 percent of neighborhoods it falls in the low income range.

similar (Figure 8). Almost half (47 percent) of them live in neighborhoods with median incomes in the extremely low or very low income categories, but 56 percent have at least one neighbor in the upper income category. However, 11 percent live in neighborhoods of the very poor: those where the highest income is in the very low or extremely low income category.

FIGURE 7 ABOUT HERE

FIGURE 8 ABOUT HERE

5.3 Mean Characteristics of Neighborhoods Conditional on Specific Characteristics of Kernel Households

The neighborhoods sampled are at different points in their dynamic evolution. Our empirical approach sets out to capture the essential feature of combined spatial and dynamic interdependence in neighborhoods. This subsection presents the *Schelling Statistics*, which measure the interdependence of kernels and their neighbors' characteristics, first for the entire population sampled and then for groups of particular interest.

5.3.1 Schelling Statistics: The Population

Our first set of the *Schelling Statistics* (conditional means and CV's) for key socioeconomic variables is given in Table 3: the variables are household income, percentage of whites, percentage of owners, age and education of head of household (reference person, according to the Codebook [Hadden and Leger (1990)]) and number of children in the household. The table compares features of the distributions of those variables across the sample of kernels for 1985 and 1993. Comparing the means gives an indication of important changes between 1985 and 1993. Mean household income for this sample increases from 1985 to 1993 though the increase is smaller than the corresponding increase in the Consumer Price Index (Urban). As a result, mean real household income decreases by 8.12 percent. The dispersion (measured by the CV) of kernels' household income registers a small decrease. There is a substantial decrease in the percentage of whites, from 72.7 percent to 65.4 percent, a much smaller decrease in the percentage of owner-occupants, from 61.2 percent to 59.9 percent, a slight increase in the average number of children, from .676 to .702, practically no change in average age, and an improvement of educational attainment of household heads, which goes up 12.4 to 12.9 years of schooling.

We obtain a first measure of the extent of mixing by computing the coefficient of variation within each neighborhood (represented by a cluster). The measures we use correspond to the benchmarks associated with the extremes of "perfect mixing" and "perfect sorting" which were discussed in section 4.2. To summarize, an increase in the mean of the CV for neighbors implied an increase in diversity within neighborhoods; an increase in the CV of the CV for neighbors implied an increase in between-neighborhood diversity.

We see that the mean (across all neighborhoods) CV of income for neighbors (within each neighborhood) decreases by 2.2 percent. The CV of the CV for neighbors (our measure of dispersion of the neighborhood CV's across all neighborhoods) decreases by 6.9 percent. These findings suggest that from 1985 to 1993 income mixing decreases both within and across urban neighborhoods in the United States. By comparing columns 2 and 3 we conclude that neighborhoods are in the average more homogeneous than the set of kernels with respect to household income, likelihood of owning and age of household head both in 1985 and 1993. This is also true for education attainment in 1993. Neighborhoods have become more homogeneous across the U.S. (as indicated by the dispersion across all neighborhoods of the neighborhood CV) with respect to income, age and presence of children. Interestingly, neighborhoods have become substantially more heterogeneous with respect to education (details are given in columns 4 and 8 of Table 3).

By looking at conditional statistics we can see how the characteristics of neighborhoods vary. Table 4 reports our second set of Schelling statistics, "conditional Schelling statistics", for the neighborhoods of kernels who have specific characteristics. These characteristics are: renters, owners, whites, nonwhites, subsidized and unsubsidized renters, subsidized and unsubsidized owners, renters and owners with children of any age, nonwhites and whites with children of any age, renters and owners with children with ages more than 10 and less than 18 years, and nonwhites and whites with children with ages more than 10 and less than 18 years.

The design of our data lends itself uniquely to this computation, which is critical for understanding key aspects of the microstructure of U.S. residential neighborhoods as they pertain to mixing in a variety of dimensions. The design of our data allows us in fact to condition precisely on characteristics of kernels, which is appropriate given that it is the kernels that are randomly chosen from among the national AHS sample.

We report in Table 4 Schelling Statistics for the same variables which were included in Table 3, and for additional variables: percentage of neighbors who are subsidized, percentage of neighbors who have very low or extremely low incomes, percentage of upper middle or high income neighbors, and percentage of neighbors who live in multi unit structures, defined as structures with more than four dwelling units. For ease of comparison, with the corresponding means and proportions in the population as a whole, we report in the column labeled "All kernels," the corresponding statistics for the kernels. We distinguish a number of categories, such as renters versus owners, and we subdivide renters further into subsidized versus unsubsidized renters, and those categories in turn into nonwhite and white household heads. We also look at households with white and nonwhite household heads without distinguishing mode of tenure. For ease of comparison the top portion of Table 4 refers to 1985 and the bottom to 1993.

5.3.2 Schelling Statistics: Tenure

Owner kernels live in neighborhoods which are more homogeneous with respect to mode of tenure. Renters have fewer neighbors who are also renters than owners do neighbors who are also owners. The typical renter has as neighbors more nonwhites(a mean of 34.1 percent) and neighbors with younger household heads (average age 45.9 years), than the typical owner (17.1 percent and 50.2 years, respectively).

5.3.3 Schelling Statistics: Race

There are noteworthy changes between 1985 and 1993 in the characteristics of subsidized renter households headed by nonwhites. For such households, the proportion of neighbors which are households headed by whites more than doubles, from 8.3 percent to 16.0 percent; the proportion of neighbors which are owners rises from 6.8 percent to 1.3 percent, and the mean education of their neighbors rises, from 9.7 to 10.8 years. Among nonwhite renters, unsubsidized households have fewer neighbors which are households with white heads (34 percent) than subsidized households (41 percent). A similar picture is suggested by the characteristics of subsidized owner and subsidized owner households headed by nonwhites.

Racial segregation is visible at the smallest neighborhood scale. Tables 3 and 4, but especially the latter, give a hitherto unavailable glimpse at segregation down at the neighborhood level. In 1985, 82.4 percent of the neighbors of white household heads were whites, while only 37.9 percent of the neighbors of nonwhite household heads are whites. There appears to be a slight reduction in this measure of racial segregation between 1985 and 1993: the corresponding figures for 1993 are 81.4 percent and 37.3 percent. By comparing households with nonwhite and white heads who rent, we note that households headed by whites have ten times more neighbors who are also whites than do nonwhites, 81 percent versus 8.3 percent.

5.3.4 Schelling Statistics: Subsidized and Unsubsidized Renters and Owners

Neighbors of subsidized renters exhibit considerably smaller mean incomes but a greater degree of income mixing (measured in terms of the CV) than those of unsubsidized renters. Income mixing, according the CV-based measure that we are using, increases among neighbors of subsidized renters from 1985 to 1993, but decreases among neighbors of unsubsidized renters. Interestingly, from 1985 to 1993, income mixing goes up among neighbors of nonwhite subsidized and unsubsidized renters, but goes down among neighbors of white subsidized and unsubsidized renters. Additional insight into income mixing is offered in subsection 5.4, where we discuss income mixing by conditioning upon characteristics of neighborhoods.

Subsidized households as a group have much lower incomes than unsubsidized ones in both 1985 and 1993. In 1985, in particular, the mean incomes for subsidized and unsubsidized households respectively are \$7075 and \$31035. Subsidized renters are 43 percent nonwhite, and nonwhites in both the subsidized and unsubsidized groups have even lower incomes, \$5140 and \$27832 respectively. The neighborhood median incomes are correspondingly low for subsidized renter and owner kernels. More than 80 percent of the neighbors of both subsidized renters and subsidized owners have very low or extremely low incomes. Fewer than 5 percent of their neighbors have incomes above HAMFI. The percentage of owners among neighbors of subsidized renters is small: 6.8 percent for nonwhites and 8 percent for whites, in 1985, and 1.3 percent and 8 percent, respectively in 1993. Subsidized renters live primarily in multi unit structures, more than two-thirds of them in structures with more than 4 units and about one-third in structures of more than twenty units in both 1985 and 1993. Subsidized renters have primarily other subsidized renters as their immediate neighbors: 86.4 percent in 1985 and 77.5 percent in 1993.

In contrast, unsubsidized renters have few subsidized renters as neighbors: 3.9 percent in 1985 and 5.5 percent in 1993. About a fourth of their neighbors are owners. They live primarily in structures with fewer dwelling units than subsidized renters: in 1985 and 1993, respectively, 26 percent and 33 percent of them live single-dwelling unit structures, and 44.8 percent and 39.5 percent in structures with more than 4 units.

Turning now to a comparison of income mixing and other neighborhood characteristics for neighbors of subsidized with those of unsubsidized owners reveals some interesting features (the sample of subsidized owners is so small that the results must be interpreted as suggestive rather than definitive). In general, the neighbors of unsubsidized owners are quite similar regardless of whether those kernels have nonwhite or white household heads, with the exception of the racial makeup of their neighborhoods. If the subsidized owner kernel is white, the proportion of neighbors who are white is 96 percent; if the kernels are nonwhite, 78 percent of their neighbors are white. The neighborhood income gap between neighbors of white and nonwhite kernels almost disappears from 1985 to 1993, from $35543 for white subsidized owners and $29628 for nonwhites, it becomes $44188 for white subsidized owners and $43400 for nonwhites.

5.3.5 Schelling Statistics: Households with Children

A key feature of the social fabric of neighborhoods is the presence of children. Children's interactions are a powerful determinant of the character to urban neighborhoods, and the presence of children in particular age groups may be a crucial factor in understanding the "epidemic" nature of antisocial activities. It is thus of particular interest to consider both income mixing and the socioeconomic structure of neighborhoods when children are present.

The neighbors of kernels with older children between 10 and 18 years old do not differ substantially from neighbors of all kernels with children. Table 4 details comparisons across neighborhoods whose kernels have children of any age and, alternatively, kernels with older children. Generally, neighbors of households with children are less likely to live in multi unit structures, which of course makes sense in view of the higher demand for space by households with children. It appears to be also true that neighbors of households with no children do not differ dramatically from those with children. The most noteworthy differences we found were between the neighbors of nonwhite-headed households with older children (aged 10 to 18 year) and neighbors of all nonwhite households. Income mixing (as measured by the CV), presence of white neighbors, and presence of owner neighbors are greater among the former than the latter, and some changes have taken place between 1985 and 1993. In particular, income mixing has increased but the presence of white neighbors has decreased. Income mixing has also increased among neighbors of nonwhites generally, as well, but so has the presence of white neighbors. This conclusion is corroborated by examining changes in the percentages of neighbors in various income categories relative to HAMFI who have incomes in the same category as themselves.

5.4 Likelihood of Income Mixing

Another way to understand income mixing in neighborhoods is to compare the *probability* that kernel households will live in neighborhoods with neighbors whose incomes are similar to, or different from theirs. This section thus analyzes the clusters data using an extension of the hypothetical models of "perfect mixing" and "perfect sorting" introduced in Section 4 of this paper.

Tables 5–6 (and Tables A.4–7 in the Appendix) are designed to highlight the extent of income mixing within U.S. neighborhoods. They provide a more detailed look at the patterns of mixing which were first introduced above in Section 5.2 These tables share the following format: in the left margin, categories of income are defined relative to HAMFI; the entries in different columns across the same row report of probabilities of various events (describing the income distribution within neighborhoods) for kernels with incomes in the range defining the row.[29] Tables A.6 and A.7 summarize these findings by means of full conditional distributions, as we see below in more detail.

Table 5 describes income mixing by showing the probability that a kernel household in each income category will live in a neighborhood with specified neighborhood income measures. The

[29] Analytically speaking, these tables report Schelling statistics, too. Unlike Table 4, which reports means for neighbors conditional on certain characteristics of the kernels, these tables report Schelling statistics which are probabilities. They are probabilities of events defined in terms of the neighborhood income distribution and conditional upon the kernel's position in the income distribution. The probabilities in the columns do not sum to one because the event defined in the columns of most of these tables are not mutually exclusive.

first column of data reproduces the first column of Tables 1 and 2 and shows the distribution of all kernel households by income category, corresponding to the probability that a kernel will be in each of those categories. We can now examine some characteristics of the distribution of income in the neighborhoods where the kernels live. In Column 2 we compare the probability that a kernel in different income groups will live in a neighborhood with an extremely low median income. Whereas in Figures 7 and 8 we were focussing on *kernels* with extremely low and very low incomes, and asking about the probability that they would live in neighborhoods with income measures in each category, here we are focussing on *neighborhoods* with specified characteristics (here a neighborhood median income of less than 30% of HAMFI) and asking, "What is the probability that a kernel in each successive income category (the rows) will live in such a neighborhood?"

We then examine the percentage of kernels in each income group who live in neighborhoods with an extremely low neighborhood median income (Column 2). The third column shows the percentage of kernels in each income category who live in neighborhoods with a neighborhood median income in the extremely high category (over 200% of HAMFI). We see that a few (6 percent) of the very high and even fewer (3 percent) of the extremely high income kernels live in the extremely low income neighborhoods. Even fewer of the very poor (1 percent) and none of the extremely poor live in the highest income neighborhoods.

By comparing with Figure 2, we see that extremely low income neighborhoods are some 13 percent of all neighborhoods, while the highest income neighborhoods are rare indeed (only 4 percent of all neighborhoods in 1985). When we look at neighborhoods where there is some mixing, we see that very low or extremely low income households are not uncommon in all but the highest income neighborhoods. The probability that a kernel will live in a neighborhood with more than one low income neighbor (the bottom quartile of the neighborhood income distribution is in the low income category) ranges from 81 or 82 percent (for kernels with extremely low or very low incomes) to 26 percent (for extremely high income kernels). [30]

A look at the likelihood of having several upper income neighbors (living in a neighborhood where the top quartile lies in the upper income range) reinforces the conclusion that the highest

[30] For more details, see the fourth column of Table 5, which shows the probability that kernels will live in neighborhoods where there is more than one neighbor (the lowest quartile of residents in a neighborhood of 10 households) with very low or extremely low incomes.

income households live apart: 79 percent of extremely high income kernels live in neighborhoods in which three fourths or more of their neighbors have upper incomes. However, a non-trivial fraction of low income households live in neighborhoods of this type (that is, with several neighbors whose incomes are in the upper income range). It shows that the probability that a household (kernel) will live in a neighborhood where the top quartile of the neighborhood income distribution lies in the upper income category (120% of HAMFI or more) ranges from 79 percent for extremely high income kernels to 15 percent (for very low income kernels) and 17 percent (for very low income kernels).

This table shows some evidence of change between 1985 and 1993, but interpreting those changes is daunting in the light of the concurrent changes in the income distribution as a whole. Does the decrease in the proportion of extremely high income kernels living in extremely high and high income neighborhoods (a decrease from 26 percent to 16 percent for neighborhoods with extremely high median incomes; a decrease from 79 percent to 73 percent for neighborhoods with a third quartile in the upper income category) indicate less mixing, or is it a statistical artifact? Similarly, does the small decrease in the proportion of extremely low and very low income kernels living in extremely low income neighborhoods imply more mixing? How do we explain the concurrent increase in the probability that extremely low income kernels will live in high income neighborhoods (an increase from 17 percent to 21 percent? We discussed above and emphasize here the problems associated with drawing conclusions about changes *over time* using measures defined in terms of HAMFI.

One measure of income mixing for low income households is access to neighborhoods where one half or three fourths of the neighbors have middle or upper incomes. In Table 6 we look at the probability that kernels in each income category will live in such neighborhoods. We then look separately at (and can compare) the probabilities for two subsets of kernels, subsidized renters and unsubsidized renters in each income category. We look first (column 1) at neighborhoods where at least 25 percent of households are in the middle or upper income groups (80 percent of HAMFI or more). A kernel household with an extremely low income has a 33 percent probability of living in such a neighborhood, compared with the 95 percent probability of living in such a neighborhood for a kernel household with an extremely high income. The same extremely low income households have a 20 percent probability of living in a neighborhood where half the households have middle or upper incomes. Virtually all those kernels with extremely low incomes who live in middle or upper income neighborhoods are either unsubsidized renters or owners. In 1985, none were subsidized renters.

Another indicator of the extent of mixing is the probability that a low income kernel household will live near neighbors with incomes higher than its own.[31] That is affected by the type of urban area a kernel lives in. The probability that a kernel household with income in each category will live in a middle or upper income neighborhood is different for central city residents, suburban residents and residents of urban (nonmetropolitan) areas. Extremely or very low income kernel households are more likely to live in a middle or upper income neighborhood if they live in a suburb of a metropolitan area (a probability of 25 percent for extremely poor kernels and 22 percent for very poor kernels); the probability is only about half as much for extremely low income and very low income kernels in central city neighborhoods (14 percent and 12 percent respectively). At the other tail of the income distribution for kernels, extremely high income households are more likely to live in a middle or upper income neighborhood if they live in a suburb (a probability of 86 percent) than if they live in a central city (64 percent) or urban nonmetropolitan neighborhoods (67 percent).

We know that single family housing is highly segregated from large multifamily structures in the U.S. housing stock. We can therefore ask how the size of the structure where a subsidized tenant lives affects the income of the neighborhood. The U.S. housing stock varies considerably by region, however. We compared relative incomes in the neighborhoods where subsidized renters live (by region), finding that dwelling type makes a difference. Subsidized renters in single family dwellings tend to live in neighborhoods with higher median incomes than subsidized renters in structures with more units. The poorest neighborhoods (those ranked lowest in terms of neighborhood median income relative to cluster HAMFI) are those where the subsidized renters live either in 2- to 4-unit structures, or in structures with 20 or more units. The small numbers of observations for subsidized kernels make analysis at this level of disaggregation both unstable and unreliable, but the findings shown in this table suggest that structure size is a dimension which warrants further exploration. For example, are subsidized tenants in 2- to 4-unit structures more likely to be renting from resident

[31] Details of this analysis are given in Table A.5.

landlords, and tenants in larger structures more likely to be in publicly owned or institutionally managed structures?

We can also compare the probability of living in a neighborhood with significant numbers of upper income households for different types of tenure and different types of structure: unsubsidized and subsidized renters, unsubsidized owners, and residents of single-unit and multi-unit structures.[32] Examining mixing in this kind of detail, we see that at each income level, kernel households who live in single family structures, and owner-occupier kernels, are more likely to live in upper income neighborhoods. The only exception is upper income renters in multifamily structures, who are no less likely to live in upper income neighborhoods than kernels with similar incomes in single family structures.

Finally, we can examine in more detail two aspects of the pattern of mixing: First, the joint distribution of kernel incomes and their neighborhood median incomes, and second, the full set of measures of the underlying neighborhood income distribution. Looking first at the joint distribution of kernel incomes and neighborhood median incomes, for example, we can compare the probability the probability that a kernel household will live in a neighborhood where their income is close to the neighborhood median, by looking at the diagonal entries in a table where the same categories are used for kernel incomes and neighborhood median income.[33] The likelihood that a kernel household lives in a very poor or extremely poor neighborhood is, not surprisingly, strongly influenced by the income of the kernel household. The kernels most likely to live in neighborhoods with medians close to theirs are the richest, followed by the poorest households (Table A.6). In 1985 kernels who have extremely low incomes (16 percent of all kernels) have a 42.7 percent probability of living in a neighborhood where the median income of its neighbors is also extremely low. The same kernel households have only an 18.3 percent probability of living in a neighborhood with median income in the middle or upper income range. At the other end of the income distribution, kernels with extremely high incomes (12 percent of all kernels) have a 64 percent probability of living in a neighborhood where the median income of their neighbors is extremely high.

In contrast, middle income households are much less likely to live in a neighborhood with a median income similar to theirs. The 16 percent of households with moderately low incomes (50%

[32]Table A.4 provides detail.

[33]Details are given in Table A.6.

to 80% of HAMFI) have only a 20 percent probability of living in a neighborhoods with a median income in the same range. Similarly the 13 percent of kernels with lower middle incomes (80% to 100% of HAMFI) have a 20.8 percent probability of living in a neighborhood with a median income in the same range as their own. Again caution is needed in interpreting results, not least because our quantiles based on HAMFI are not of equal sizes.

A comparison of the tables for 1985 and 1993 suggests that higher income neighborhoods may be becoming more accessible to the poorest households, and that the highest income households are slightly less likely to live in exclusively high income neighborhoods, but this conclusion is tempered by our previous caveats and the observation that the categories for both households and neighborhoods are defined in terms of HAMFI, and that the ratio of HAMFI to median and mean incomes is higher in 1993 than in 1985.

Turning to the more detailed analysis of distribution of neighborhood incomes, Table A.7 shows the probability that income distribution measures for a *neighborhood* will fall in each quantile, defined in terms of HAMFI. Incomes are dispersed in most U.S. neighborhoods, in the sense that the extremes of the income distribution are represented in the majority of neighborhoods. In 1985, the minimum income (that is, the income of the poorest household) falls in the extremely low range (0 to 30% of HAMFI) in 66.1 percent of all neighborhoods. The highest income households are also widely dispersed: in the same year, the income of the richest neighbor is extremely high (over 200% of HAMFI) in 36.5 percent of neighborhoods, and there is at least one very high or extremely high income neighbor in 58.5 percent of all neighborhoods. The same table also shows some neighborhoods which look more like "perfect sorting": neighborhoods made up of concentrations of households all of whom have very similar incomes. We see such neighborhoods at both extremes of the income distribution. In 12.1 percent of neighborhoods, three fourths of all households have extremely low or very low incomes. In 4.9 percent of neighborhoods, the *richest* household has a very low or extremely low income. At the other extreme of the income distribution, 10.8 percent of neighborhoods have a very high or extremely high median income. In a very few (3 percent) of neighborhoods in our sample, segregation is so great that only the poorest 25 percent (the bottom quartile of the income distribution) have incomes below the very high income range. In all but a few of those high income neighborhoods, we find some low income households. In 96 percent of all neighborhoods at least one household has a low, very low or extremely low income. Rather than interpret this as unambiguous evidence of mixing, however, we should note that some at least of the low income households in higher income neighborhoods have higher permanent incomes, but their current in comes may reflect either short term fluctuations in income (for household heads recently downsized or laid off) or are living on accumulated assets (retirees or rentiers).

6 Comparison with Evidence of Income Mixing in Large Scale Urban Neighborhoods

Communities display degrees of both homogeneity and heterogeneity in many dimensions and at many scales. We know that even where there is mixing by race in a large geographical area, there may still be segregation at the local (immediate neighborhood) level [Massey and Denton (1993)]. Massey and Hajnal (1955) show how a long-term trend away from macro-level and towards micro-level segregation has helped segregation patterns to evolve to minimize, in particular, white contact with blacks. As we have noted, however, we know much less about both static and dynamic patterns of segregation by income. The micro level of detail provided by the AHS clusters allows us to gauge the extent of such segregation or mixing, by race or by income level at a much smaller scale.

Most of what we know about the extent of segregation or mixing in U.S. cities (and much more is known about U.S. cities than about urban areas in any other country) has been based on aggregate data on neighborhoods defined by census tracts. In contrast, the neighborhoods considered here are defined as the ten neighbor households closest to the kernel. This is a neighborhood defined with a very small grain, at the level of the building or street. Whereas census tract "neighborhoods" correspond more closely to the notion of a school district, the neighborhood clusters "neighborhoods" correspond to the households who are seen and with whom households interact daily, on stairs, in elevators, at the mailbox or in the street or playground.

As Galster and Killen point out "there is no single geographic scale over which the metropolitan opportunity structure varies. On certain dimensions it may be relatively invariant across an entire metropolitan area, varying only from one metropolitan area to the next. More likely, certain dimensions vary across municipal jurisdictions and others vary across census tracts within municipalities... . Still others may vary at even smaller scales. (p. 21)" This section makes some comparisons between findings of the existing literature on income segregation and mixing and the characteristics of neighborhoods as they are defined by the AHS neighborhood clusters data set.

Empirical studies of mixing using aggregate data make extensive use of indices of segregation and dissimilarity. These aggregate measures of the degree of segregation, mixing or concentration of specific groups in space are standard tools. The measures depend on the availability of observations for *all* the spatial areas which make up the city or metropolitan area. The clusters in contrast sample households and their neighbors, and represent only a tiny fraction of each metropolitan area. What we can see from the clusters is how the characteristics of neighbors vary, conditional on the kernel's characteristics.

Evidence of the relative extent of segregation by race and by income is mixed. One source of variation is the definitions of poverty and race which are used. Abramson, Tobin and VanderGoot (1995) using the U.S. census definition of poverty, conclude that in 1990 "compared with the segregation of the poor, the segregation of blacks was much greater in the 100 largest U.S. metropolitan areas in 1990 ... the dissimilarity for blacks was 60.6, compared with the 36.1 dissimilarity for the poor..."(p. 51). The use of averages conceals wide variation between metropolitan areas: the same authors found that the index of dissimilarity for the poor in the 100 largest metropolitan areas in 1990 ranged from 19.6 to 55.1. Studies using more narrowly defined income classes have (not surprisingly) found more segregation. Abramson *et al.* contrast their results with Massey and Eggers estimate of a dissimilarity index of 47.0 for "poor" versus "affluent" families in 60 metropolitan areas in 1980 and "an average dissimilarity across the four income classes of 29.4." (p.53).

White (1987) found that the poor (defined as households or individuals below the poverty threshold) are significantly segregated - with a dissimilarity index of 36%, meaning that on average about thirty six percent of poor families would have to move to be evenly distributed throughout the SMSA. The corresponding measure of dissimilarity for race (black) was 67.6 percent ; for the college educated it was 32.6 percent. Multi-family housing was almost as concentrated as race (an index of 52.8 for multi-family housing and 43.3 for single family dwellings. Home ownership had an index of 38.6. However, he also pointed out that poverty dissimilarity disaggregated by race was much less than for the total population (29.2 on average for whites and 32.4 for blacks), implying that segregation by race and the correlation of race and income have confounding effects.

We can compare White's ranking of his index of dissimilarity, ranking variables by the degree of dissimilarity which they display on average in his sample of cities. When we rank household characteristics by the degree of dispersion exhibited by our neighborhoods (using the CV as the measure of dispersion) we find a pattern comparable but not identical to that found by White, with the exception of the number of children per household (which has a coefficient of variation of 1.59 in 1985 and 1.56 in 1993) and (more surprisingly) education of the household head (which has a coefficient of variation of .294 in 1985 and .262 in 1993). In our data, race has a lower coefficient of variation (.62) than income (.87) or tenure (.80) in 1985, although the difference narrows from 1985 to 1993 when the corresponding values are income (.853) tenure (.789) and race (.731). Both White and our data show much less dispersion in age between neighborhoods. In our data, as in White's, housing types are very highly segregated, with a coefficient of variation of.88 for single family housing and 1.27 for multi- family structures with more than 5 units.

7 Conclusion

The neighborhoods sampled are at different points in their dynamic evolution. Because the American Housing Survey neighborhood clusters data were collected for several periods, they can and should be used to study the dynamics of mixing – the evolution of households' and dwellings' circumstances, conditional upon specific characteristics of their neighborhoods – in much more detail. Our empirical approach, conditioning upon neighbors' characteristics as well as their decisions, set out to capture the essential features of spatial interaction in neighborhoods.

What our two snapshots of neighborhoods are not able to capture is the phenomenon of combined spatial and dynamic interdependence. Individuals do care about their neighbors' characteristics, and households' location decisions interact in extremely complex ways to produce the outcomes of mixing and segregation which we observe. Just as households' incomes and their tastes affect residential patterns, we also know that individuals' choices and their opportunity sets are affected by the characteristics of their neighbors, or by changes in their immediate neighborhood. Residential choices are affected, as well as choices about education, employment, teenage pregnancy, or drug use. Further research on the dynamics of mixing and segregation using these data will be invaluable.

Further research on the dynamics of segregation and mixing, however, will demand careful definition of quantiles of the distribution of incomes which are consistent over time. It is clearly appropriate to take account of the impact of household size on the standard of living represented by a given household income. With steadily declining household sizes, declines in household income need to be interpreted with care. At the same time, it is imperative that we take account of changes the price structure, while avoiding the distortions which are potentially introduced by any simple measure of the effect of inflation on real standards of living. The research presented here makes it clear that variations in changes in the cost of living from city to city are significant, and that they cannot be ignored in looking at the dynamics of housing markets. An obvious explanation is that housing costs make up a substantial component in any cost of living index. But housing costs (and their rate of change) probably vary more from city to city than any other component in the cost of living.

Future work (perhaps using quintiles rather than quantiles defined with respect to HAMFI) would allow us to examine the dynamic impact of institutional and individual forces at work in facilitating or reducing segregation, such as the presence or absence of households who are recipients of housing subsidies, the age of household heads, the presence or absence of children. Models of individual choice, for example, could be used to analyze the characteristics of movers out (households who were present in 1985 and had moved before 1989 or 1993) and (more important) of their neighborhoods. Similarly, using individual choice models, we could look at interactions between neighborhood features and characteristics of movers in (households present in 1985 and not in 1989 or 1993).

A provocative outcome of this study is that it identifies neighborhoods where the poor are isolated - those where the third quartile or even the maximum neighborhood income falls in the very low income or extremely low income categories. But it also suggests that households do not care so much about their neighbors' incomes that they sort themselves out in a pattern remotely resembling the "perfect sorting" model we used as a benchmark. We found one or more low income households and one or more high income households in the great majority of the kernels' neighborhoods.

A next step is to explore further the dynamics of this pattern. How many of the neighborhood outliers – households with very low or very high incomes in otherwise middle income neighborhoods, for example – have incomes which are only temporarily different from most of their neighborhoods. Are the neighborhoods apparently "mixed" because some residents have experienced temporary shocks which caused their income to be unusually high or low? Are the neighborhoods apparently "mixed" because some households have unusually high (or low) tastes for housing, or for neighborhood characteristics, and thus choose to live in neighborhoods where they spend much more (or less) on housing than their neighbors? Are the neighborhoods mixed because moving is costly, and therefore many, perhaps most households are not in equilibrium? In that case the neighborhood "outliers" will be the first to move. Or have some "outliers" chosen to stay because they have strong ties to their neighborhood. Each explanation has different policy implications.

BIBLIOGRAPHY

Abramson, Alan J., Mitchell S. Tobin and Matthew R. VanderGoot "The Segregation of the Poor in U.S. Metropolitan Areas" *Housing Policy Debate* Vol.6, Issue 1, 1995, pp. 45-72.

Bradbury, Katharine L. (1996), "The Growing Inequality of Family Incomes: Changing Families and Changing Wages," *New England Economic Review,* July/August, 55-82.

Bratt, Rachel (1989), *Rebuilding a Low-Income Housing Policy,* Temple University Press, Philadelphia, PA.

Crane, Jonathan (1991), "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing," *American Journal of Sociology,* 96, 5, 1226-1259.

de Bartolomé, Charles and Stuart S. Rosenthal (1996a), "Opposites Attract: The Effect of The Federal Income Tax Code on Community Composition," University of British Columbia, mimeo.

de Bartolomé, Charles A. M. and Stuart S. Rosenthal (1996), "Property Tax Capitalization in a Model With Tax Deferred Assets and Standard Deductions," working paper, University of British Columbia.

Ellickson, Bryan (1979), "Local Public Goods and the Market for Neighborhoods," 263-292, D. Segal, ed., *The Economics of Neighborhood,* Academic Press, New York.

Esteban, Joan-Maria, and Debraj Ray (1994), "On The Measurement of Polarization," *Econometrica,* 62, 4, 819-851.

Gabriel, Stuart A., and Stuart S. Rosenthal (1996a), " Commute Times, Neighborhood Effects, and Earnings: An Analysis of Compensating Differentials and Racial Discrimination," *Journal of Urban Economics,* 40, July, 61-82.

Gabriel, Stuart A., and Stuart S. Rosenthal (1996b), " Location and the Effect of Demographic Traits on Earnings: The Tradeoff Between Omitted Variable and Simultaneity Bias," working paper, University of British Columbia.

Galster, George C. (1987), *Homeowners and Neighborhood Reinvestment,* Duke University Press, Durham, North Carolina.

Galster, George C., and Sean P. Killen (1995), "The Geography of Metropolitan Opportunity: a Reconnaissance and Conceptual Framework"*Housing Policy Debate* Vol.6, Issue 1, 7-43.

Goodman, Allen C. (1989), "Topics in Empirical Urban Housing Research," in Muth, Richard F., and Allen C. Goodman (1989), *The Economics of Housing Markets,* Harwood Academic Publishers, Chur, Switzerland, 54-81,

Gladwell, Malcolm (1996), "The Tipping Point," *The New Yorker,* June 3, 32-38.

Hadden, Louise, and Mireille Leger (1990), *Codebook for the American Housing Survey,* Abt Associates, Cambridge, MA.

Hoyt, William H., and Stuart S. Rosenthal (1995), " Household Location and Tiebout: Do Families Sort According to Preferences for Locational Amenities?" *Journal of Urban Economics,* forthcoming.

Ioannides, Yannis M. (1996) " Residential Neighborhood Effects and Inequality, " Tufts University, July (mimeo).

Jencks, Christopher, and Susan E. Mayer (1990), "The Social Consequences of Growing Up in a Poor Neighborhood," in M. G. H. McGeary and L. E. Lynn, eds., *Inner City Poverty in America,* National Academy Press, Washington, DC.

Khadduri, Jill, and Kathryn P. Nelson (1992), "Targeting Housing Assistance," *Journal of Policy Analysis and Management,* 11, 1, 21-41.

Massey, Douglas S., and Nancy A. Denton (1988b), "The Dimensions of Residential Segregation," *Social Forces,* 67, 2, 281-315.

Massey, Douglas S., and Nancy A. Denton (1993) *American Apartheid: Segregation and The Making of The Underclass,* Harvard University Press, Cambridge, MA.

Massey, Douglas S., and Zoltan L. Hajnal (1995), "The Changing Geographic Structure of Black-White Segregation in The United States," *Social Science Quarterly,* 76, 3, 527-542.

Mayer, Christopher J. (1996), "Does Location Matter?" *New England Economic Review,* May/June, 26-40.

Muth, Richard F. (1969), *Cities and Housing,* University of Chicago Press, Chicago, IL.

Schelling, Thomas. C., (1971), "Dynamic Models of Segregation," *Journal of Mathematical Sociology,* 1, 143-186.

Schelling, Thomas C., (1978), *Micromotives and Macrobehavior,* Norton, New York.

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

U.S. Bureau of the Census (1985; 1993) *American Housing Survey of the United States,* Current Housing Reports H150/85, H150/93, U.S. Department of Commerce, Washington, DC.

U.S. Bureau of the Census (1987; 1995), *Statistical Abstract of the United States,* U.S. Department of Commerce, Washington, DC.

U.S. Bureau of the Census (1996), *American Housing Survey,* National Core and Supplement File, CD-ROM (recordable) for 1985, 1987 and 1989, and 1989, 1991, and 1993, U.S. Department of Commerce, Washington, DC.

U.S. Department of Housing and Urban Development (1996), *Rental Housing Assistance at a Crossroads: A Report to Congress on Worst Case Housing Needs,* Office of Policy Development and Research, March.

Vandell, Kerry D. (1995), " Market Factors Affecting Spatial Heterogeneity among Urban Neighborhoods," *Housing Policy Debate,* 6, 1, 103-139.

White, Michael J. (1987), *American Neighborhoods and Residential Differentiation,* Russell Sage Foundation, New York.

Figures

Figure 1
Income Categories Used in the Analysis

Income as % of HAMFI	Terms Used in this Study	Corresponding HUD - Defined Income Category	
	Low income households: *48%*		
0-30%	Extremely low income *17%*	Extremely low income	
		Very low income	
30-50%	Very low income *13%*		
			Low income
50-80%	Moderately Low income *18%*		
	Middle income households: *21%*		
80-100%	Lower middle income *12%*		
		Middle Income	
100-120%	Upper middle income *9%*		
	High income households: *31%*		
120-150%	Moderately High income *10%*		
			Upper Income
150-200%	Very high income *10%*		
200% -	Extremely high income *11%*		

Figures in italics are the proportion of kernel households in each category in 1993

Figure 2
Incomes of Households and Median Incomes of Neighborhoods

	1985		1993	
Income as % of HAMFI	**Household** Income: % of Households	**Neighborhood** Median income: % of Clusters	**Household** Income: % of Households	**Neighborhood** Median Incom % of Clusters
0-30%	16	13	17	13
30-50%	13	12	13	16
50-80%	16	27	18	27
80-100%	13	18	12	16
100-120%	9	10	9	12
120-150%	10	11	10	9
150-200%	11	5	10	5
200% -	12	4	11	2
Total	100	100	100	100

Figure 3

Measures of Income Distribution and Dispersion: Coefficients of Variation for Households and Neighborhoods

	1985	1993
Household Income (Kernels)	**$**	**$**
Mean	29755	36712
Standard Deviation:	25937	31320
Coefficient of Variation:	87.2	85.3

NeighborhoodMedian Income: The distribution of Coefficients of Variation

	1985	1993
Mean of Nbhd CV	60.4	62.7
Minimum	7.5	7.6
Q1	46.1	48.7
Q2(Median)	58.2	60.3
Q3	71.2	73.7
Maximum	192.7	179.5

Figure 4
Hypothetical Distribution of Incomes of Neighbors : "Perfect Mixing"

Income distribution of Neighbors for all kernels
If all neighborhoods were identical (all households live in neighborhoods with an income distribution identical to the population income distribution)

Measure as % of HAMFI	Neighborhood Min Income: % of Clusters	Neighborhood Median Income: % of Clusters	Neighborhood Max. Income: % of Clusters
0-50%	100%	0	0
50-120%	0	100%	0
>120%	0	0	100%
Total	100	100	100

Figure 5
Hypothetical Distribution of Incomes of Neighbors: "Perfect Sorting"

Note: Income distribution of the population is assumed to be that of the U.S. in 1985

Income distribution of Neighbors for all kernels if all households were perfectly sorted (all residents of a neighborhood are identical to each other)

Income as % of HAMFI	Neighborhood Min Income: % of Clusters	Neighborhood Median Income: %of Clusters	Neighborhood Max. Income: % of Clusters
0-50%	29	29	29
50-120%	38	38	38
>120%	33	33	33
Total	100	100	100

Figure 6
Actual Distribution of Incomes of Neighbors: All Kernels 1985

Actual Income distribution of Neighbors

Measure as % of HAMFI	Neighborhood Min Income: % of Clusters	Neighborhood Q1 Income: % of Clusters	Neighborhood Q3 Income: %of Clusters	Neighborhood Max. Income: % of Clusters
0-50%	84.1	52.7	12.1	4.9
50-120%	14.2	40.1	45.5	22.7
120%-	1.6	7.0	42.4	72.5
Total	100	100	100	100

Figure 7
Distribution of Incomes of Neighbors:Extremely Low Income Kernels 1985

Summary Distribution of Incomes of Neighbors: Extremely Low Income Kernels 1985

Income distribution of neighbors:

For kernels with income <.3HAMFI in 1985 (N=82)

Income as % of HAMFI	Neighborhood Min Income: % of Clusters	Neighborhood Median Income: %of Clusters	Neighborhood Max. Income: % of Clusters
0-50%	96	59	14
50-120%	4	36	39
>120%	0	6	46
Total	100	100	100

Detailed Distribution of Incomes of Neighbors: Extremely Low Income Kernels 1985

Income as a % of HAMFI	Neighbors Min Income: % of Clusters	Q1 Income:	Median	Q3 Income:	Max. Income:
0-30%	89	69	39	22	7
30-50%	7	11	20	15	7
50-80%	4	16	27	30	12
80-100%	0	3	4	7	21
100-120%	0	0	5	9	6
120-150%	0	1	6	12	13
150-200%	0	0	0	3	17
200%	0	0	0	3	16
Total	100	100	100	100	100

Figure 8
Distribution of Incomes of Neighbors: Very Low Income Kernels 1985

Income distribution: 1985
For kernels with income between .3 HAMFI and .5 HAMFI in 1985 (N=73)

Summary Distribution of Incomes of Neighbors: Very Low Income Kernels 1985

Income as % of HAMFI	Neighborhood Min Income: % of Clusters	Neighborhood Median Income: %of Clusters	Neighborhood Max. Income: % of Clusters
0-50%	99	47	11
50-120%	1	45	33
>120%	0	8	56
Total	100	100	100

Detailed Distribution of Incomes of Neighbors: Very Low Income Kernels 1985

Income as % of HAMFI	Neighbors Min Income: % of Clusters	Q1 Income:	Median	Q3 Income:	Max. Income:
0-30%	88	45	21	8	1
30-50%	11	37	26	15	10
50-80%	1	11	36	29	11
80-100%	0	5	8	21	11
100-120%	0	1	1	12	11
120-150%	0	0	5	8	18
150-200%	0	0	1	4	19
200%	0	0	1	3	19
Total	100	100		100	100

Tables

Table 1:

Income Distribution Relative to HUD-adjusted Median Family Income, by Regions, 1985 and 1993

Sample: Kernels and Neighbors

Year	1985					1993				
Regions	All SMSAs	Mid West	North East	South	West	All SMSAs	Mid West	North East	South	West
Mean HAMFI ($)	30845	31001	34267	27608	32381	41365	41574	45954	37023	43425
0%–30%	16	17	19	15	14	17	18	18	16	16
30%–50%	12	13	12	12	12	13	15	12	13	15
50%–80%	16	18	16	15	17	18	19	17	17	16
80%–100%	13	14	14	10	13	12	12	13	11	14
100%–120%	9	10	9	9	10	9	9	8	9	8
120%–150%	11	9	10	12	11	10	11	10	9	10
150%–200%	11	10	11	12	11	10	9	10	11	10
200%–	12	9	10	15	11	11	8	11	13	11
Total	100	100	100	100	100	100	100	100	100	100

Table 2:

Income Distribution Relative to HUD-adjusted Median Family Income, by Central City, Nonmetro Area and Suburban Areas of Central Cities

Sample: Kernels and Neighbors

Year	1985				1993			
Regions	All SMSAs	Central City	Urban Nonmetro	Suburban	All SMSAs	Central City	Urban Nonmetro	Suburban
Observations	6215	2946	663	2606	9207	4126	976	4105
Total (%)	100	47	11	42	100	45	11	45
Mean HAMFI ($)	30845	29879	24197	32299	41365	40958	32862	42798
Distribution of Household Income as % of HAMFI								
0%–30%	16	21	16	10	17	22	14	12
30%–50%	12	14	13	10	13	15	14	12
50%–80%	16	17	17	16	18	18	19	17
80%–100%	13	11	9	15	12	12	9	14
100%–120%	9	8	9	11	9	8	10	9
120%–150%	10	10	9	12	10	9	11	11
150%–200%	11	9	11	13	10	8	9	12
200%–	12	10	15	13	11	9	14	12
Total	100	100	100	100	100	100	100	100

Table 3:

Mean Characteristics of Kernel Households and Neighborhoods

("Schelling Statistics")

Column 1 reports means for the respective variables for kernel households, for 1985.

Column 2 reports the coefficient of variation, defined as sample standard deviation divided by sample mean, among all kernel households.

Column 3 reports the mean of the neighborhood coefficient of variation, to be referred to as neighborhood CV, which is computed by associating with each kernel the coefficient of variation of respective variable among its neighbors and then by taking the mean over all kernels.

Column 4 reports across all neighborhoods the coefficient of variation of the neighborhood CV, which is computed by associating with each kernel the coefficient of variation for the respective variable among its neighbors and then taking the coefficient of variation over all kernels.

Columns 5–8 repeat the same measures for 1993.

Year	1985				1993			
Column	1	2	3	4	5	6	7	8
Statistic	Mean	CV	Mean	CV	Mean	CV	Mean	CV
Sample	Kernels	Kernels	Nei/hoods	Nei/hoods	Kernels	Kernels	Nei/hoods	Nei/hoods
Observations	680	680	680	680	960	960	960	960
Income	29755	.872	.605	.361	36712	.853	.627	.336
White (=1)	.727	.615	.619	1.15	.654	.728	.731	1.03
Owner (=1)	.612	.797	.764	1.08	.599	.818	.789	1.06
Age (years)	49.1	.368	.306	.328	49.33	.358	.308	.307
Education (years)	12.4	.294	.306	.328	12.9	.262	.205	.585
Children	.676	1.59	1.604	.422	.702	1.56	1.680	.403

Table 4:

Mean Characteristics of Neighbors Conditional on Characteristics of Kernel Households

(Conditional "Schelling Statistics")

Variables	Kernels	Neighbors of			
	Only	Renter Kernels	Owner Kernels	White Kernels	Nonwhite Kernels
Year		1985 Characteristics of Neighbors			
Income	29755	21917	34551	31157	22227
Inc (CV)	.872	.960	.759	.825	.894
White (=1)	.838	.774	.874	.94	.297
Owner (=1)	.616	.228	.840	.639	.453
Age (yrs)	49.13	45.85	50.24	49.03	46.60
CV	.368	.408	.328	.357	.361
Education (yrs)	12.39	11.98	12.86	12.71	11.41
CV	.294	.308	.260	.266	.336
Children	.68	.64	.73	.67	.88
CV	1.59	1.72	1.52	1.61	1.42
Subsidized (%)	5.2	13.2	0.6	3.7	14.5
% below 50 % HAMFI	27.2	43.3	19.2	25.0	45.5
% above 100% HAMFI	42.7	28.1	52.1	45.7	29.2
% in multiunit	20.5	47.9	3.8	17.5	33.4
Obs	581	2206	3591	4710	914
Year		1993 Characteristics of Neighbors			
Income	36712	29090	43903	39802	30809
Inc (CV)	.853	.918	.769	.814	.957
White (=1)	.817	.753	.869	.930	.322
Owner (=1)	.599	.261	.834	.642	.458
Age (years)	49.33	46.14	51.81	50.00	47.39
CV	.358	.396	.325	.354	.360
Education (yrs)	12.95	12.60	13.25	13.13	12.39
CV	.262	.271	.233	.242	.271
Children	.70	.68	.65	.63	.85
CV	1.56	1.70	1.62	1.67	1.52
Subsidized	5.5	13.5	.8	3.7	15.7
% below 50 % HAMFI	30.3	43.0	20.9	26.3	43.8
% above 100% HAMFI	39.7	27.7	49.2	43.5	29.2
% in multiunit	20.1	44.8	4.6	17.7	31.2
Obs	812	3201	5084	6516	1380

Table 4: Continued

(Conditional "Schelling Statistics")

Variables	Subsidized Renters		Unsubsidized Renters		Subsidized Owners		Unsubsidized Owners	
	NonWhites	Whites	NonWhites	Whites	NonWhites	Whites	NonWhites	Whites
Year	1985 Characteristics of Neighbors							
Income	7846	9200	17499	24724	42895	33226	29628	35543
Inc (CV)	.983	1.102	.823	.872	.48	.71	.757	.758
White (=1)	.0930	.807	.414	.919	.472	.92	.22	.96
Owner (=1)	.698	.077	.198	.242	.962	.81	.799	.86
Age (yrs)	46.87	57.99	43.83	44.71	45.9	44.5	49.5	50.8
CV	.370	.346	.399	.412	.291	.35	.32	.32
Education (yrs)	9.65	10.42	11.36	12.40	13.7	12.5	12.0	13.0
CV	.327	.339	.325	.290	.213	.251	.343	.246
Children	1.42	.558	.776	.559	.943	.97	.784	.71
CV	1.08	1.88	1.54	1.81	1.03	1.18	1.50	1.54
Subsidized	86.0	75.0	3.9	2.0	0.0	1.4	1.7	2.003
% below 50 % HAMFI	88.4	84.0	52.9	32.8	1.9	16.4	28.2	17.9
% above 100% HAMFI	3.1	4.4	21.1	33.4	60.4	52.1	43.1	54.0
% in multiunit	65.1	70.7	52.6	42.6	0.0	2.1	5.5	2.8
Obs	129	181	384	1426	53	140	348	2952
Year	1993 Characteristics of Neighbors							
Income	8647	16924	25624	33341	34603	45510	43401	44188
Inc (CV)	1.08	1.00	.931	.842	.77	.62	.78	.77
White (=1)	.158	.838	.338	.904	.493	.96	.330	.947
Owner (=1)	.005	.076	.228	.303	.841	.797	.804	.841
Age (years)	43.16	51.34	45.13	45.70	49.1	47.3	50.9	52.2
CV	.410	.408	.389	.391	.32	.32	.31	.33
Education (yrs)	10.91	11.54	12.34	12.91	12.1	13.5	13.0	13.3
CV	.271	.279	.267	.269	.229	.235	.267	.224
Children	1.26	.81	.82	.606	.65	.74	.78	.63
CV	1.16	1.67	1.55	1.77	1.91	1.36	1.61	1.62
Subsidized	87.4	60.0	7.3	3.0	1.4	0.5	1.6	0.8
% below 50 % HAMFI	88.4	68.5	48.7	34.7	31.9	13.5	25.4	20.0
% above 100% HAMFI	3.2	11.1	21.6	33.6	36.2	58.0	45.6	49.7
% in multiunit	74.7	71.5	43.1	39.5	2.9	0	5.9	4.4
Obs	190	235	536	2049	69	207	555	4005

Table 4: Continued

(Conditional "Schelling Statistics")

Variables	All Children		All Children		Children, 10-17		Children, 10-17	
	Renters	Owners	NonWhites	Whites	Renters	Owners	NonWhites	Whites
Year	1985 Characteristics of Neighbors							
Income	20359	37191	23158	33938	20609	36303	22974	33545
Inc (CV)	.87	.734	.900	.787	.887	.734	.959	.763
White (=1)	.67	.85	.28	.93	.66	.83	.30	.95
Owner (=1)	.31	.86	.51	.72	.28	.85	.49	.70
Age (yrs)	44.1	48.8	45.9	47.6	42.4	48.5	45.2	46.7
CV	.371	.320	.348	.337	.361	.313	.344	.332
Education (yrs)	11.2	12.94	11.4	12.6	10.9	12.7	11.1	12.5
CV	.346	.256	.335	.277	.380	.259	.357	.287
Children	1.10	.89	1.13	.91	1.28	.97	1.30	1.00
CV	1.18	1.33	1.20	1.30	1.11	1.21	1.08	1.23
Subsidized	16.4	0.4	14.9	3.2	16.7	0.5	16.2	2.6
% below 50 % HAMFI	48.1	17.3	45.5	22.5	49.8	17.7	49.0	21.8
% above 100% HAMFI	24.4	54.9	29.8	48.9	24.6	53.1	25.1	49.6
% in multiunit	29.5	1.6	23.3	7.4	37.0	0.1	27.0	8.3
Obs	874	1726	580	2020	468	819	359	928
Year	1993 Characteristics of Neighbors							
Income	28760	47121	32298	42576	28813	47612	30845	43233
Inc (CV)	.953	.745	.944	.802	.987	.749	.972	.810
White (=1)	.67	.84	.28	.92	.62	.83	.25	.91
Owner (=1)	.30	.86	.54	.69	.28	.87	.49	.69
Age (years)	44.7	50.3	46.8	48.6	44.4	50.0	44.9	48.8
CV	.378	.321	.352	.343	.380	.319	.357	.340
Education (yrs)	11.9	13.4	12.3	13.0	11.7	13.5	12.2	13.0
CV	.291	.227	.263	.253	.310	.231	.264	.268
Children	1.09	.79	1.14	.93	1.31	.94	1.34	1.00
CV	1.23	1.39	1.26	1.27	1.12	1.25	1.11	1.23
Subsidized	16.2	0.9	16.6	3.7	22.2	1.2	24.8	4.5
% below 50 % HAMFI	47.7	17.8	42.8	25.0	49.4	18.0	47.4	24.9
% above 100% HAMFI	24.9	52.4	30.1	45.6	24.6	51.2	26.9	45.2
% in multiunit	31.5	1.5	19.1	10.9	33.9	0	24.9	9.7
Obs	1380	2309	834	2855	658	1021	405	1274

Table 5:

Incomes of Kernels and The Neighborhoods They Live in

The kernel's income category is shown in the left margin. Successive rows of the table show the probability that a kernel household that income category lives in a neighborhood with the characteristics specified at the top of the columns. Different columns specify different neighborhood characteristics (alternative ways of describing neighborhood income mixing). Column 2 gives the probability that a kernel household in each income group will live in a neighborhood with extremely low median income. Column 3 gives the probability that a kernel household in each income group will live in a neighborhood with extremely high median income. Column 4 gives the same probability for neighborhoods where the lowest quartile of neighbors have very low or extremely low incomes, and Column 5 gives the same probability for neighborhoods where the top quartile of neighbors have very high or extremely high incomes. Note: Numbers in the rows do not sum up to one (even though they represent probabilities) because the categories (columns) are not mutually exclusive.

Year	1985					1993				
Income as % of HAMFI	All Kernels	Probability				All Kernels	Probability			
		$Q_{50} < .3H$	$Q_{50} > 2H$	$Q_{25} < .5H$	$Q_{75} > 1.2H$		$Q_{50} < .3H$	$Q_{50} > 2H$	$Q_{25} < .5H$	$Q_{75} > 1.2H$
	1	2	3	4	5	6	7	8	9	10
0–30%	16	39	0	81	17	17	37	1	85	21
30%–50%	12	21	1	82	15	13	18	1	75	29
50%–80%	16	13	2	63	31	17	13	1	66	30
80%–100%	13	5	1	43	51	13	7	0	63	40
100%–120%	9	8.	2	39	51	10	4	0	47	54
120%–150%	11	0	2	38	52	10	1	3	39	56
150%–200%	11	6	5	31	61	10	2	2	33	72
200%–	12	3	26	26	79	10	3	15	31	73

Table 6:
Incomes and Household Characteristics of Kernels in Middle and High Income Neighborhoods

This table shows the probability that a kernel household with selected household or dwelling structure characteristics will live in a middle or high income neighborhood (neighborhoods with median neighborhood income of 80% or more of the HUD-Adjusted Median Family Income for the metropolitan area).

The kernel's income category is shown in the left margin. Columns 1–5 report the probability that a kernel household in each of those income categories in 1985 lives in a neighborhood with the characteristic or characteristics specified at the top of the column; columns 6–10 repeat for 1993. Column 1 reports the probability that kernel households in each income category live in neighborhoods in which at least the richest one-half of their neighbors had middle or high incomes. In column 2 we consider neighborhoods in which households in the top quartile have incomes in the middle or high income category. The probability that a household with extremely low income will live in a neighborhood where the richest quarter of neighbors have middle or high incomes was 33 percent in 1985. Successive rows report the probability of the same event for kernel households with very low income, low income, etc. Column 3 reports the probability that kernel households who are unsubsidized renters in each income category live in neighborhoods in which at least the richest one-half of their neighbors had middle or high incomes. Column 4 reports the probability that kernels who are subsidized renters in each income category live in a neighborhood in which at least 50 percent of the neighbors had middle or high incomes (the median income of the neighborhood is 80% of HAMFI). Column 5 reports the probability that kernel households who are owners in each income category live in neighborhoods in which at least the richest one-half of their neighbors had middle or high incomes.

Note: Numbers in the rows do not sum up to one (even though they represent probabilities) because the categories (columns) are not mutually exclusive.

Year	1985					1993				
	Probability					Probability				
Income as % of HAMFI	Q_{50} $> .8H$	Q_{75} $> .8H$	Q_{50} $> .8H$ Unsubs/d Renter	Q_{50} $> .8H$ Subsidized Renter	Q_{50} $> .8H$ Owners	Q_{50} $> .8H$	Q_{75} $> .8H$	Q_{50} $> .8H$ Unsubs/d Renter	Q_{50} $> .8H$ Subsidized Renter	Q_{50} $> .8H$ Owners
	1	2	3	4	5	6	7	8	9	10
0–30%	15	33	20	0	22	18	39	13	5	37
30%–50%	18	48	14	0	22	28	57	18	4	36
50%–80%	36	68	18	0	55	27	63	19	0	33
80%–100%	60	84	31	.	75	49	75	34	0	59
100%–120%	56	84	18	0	65	58	81	46	.	64
120%–150%	70	88	43	.	77	70	91	42	0	80
150%–200%	77	89	61	.	83	70	93	41	.	79
200%–	74	95	36	.	82	73	94	75	.	72

Table 7:

Incomes in Neighborhoods of Multifamily Structures

Means are reported for a categorical variable defined for neighborhoods with kernels who are unsubsidized renters, as follows. For each neighborhood the categorical variable is assigned a value of 1, 2, and 3, depending upon whether the ratio of median cluster income to cluster HAMFI (calculated for the cluster-average household size) for the neighborhood belongs to the low, medium, and upper one-third, respectively.

Year	1985	1985	1985	1985	1993	1993	1993	1993
Region	Midwest	Northeast	South	West	Midwest	Northeast	South	West
1 unit	2.00	2.00	1.94	1.89	1.76	2.18	1.84	1.93
2–4 units	1.27	1.81	1.60	1.00	1.55	1.63	1.20	1.40
5–19 units	1.67	1.50	1.79	1.83	1.80	1.86	1.30	1.67
more than 20 units	1.60	1.77	1.75	1.80	1.20	1.75	2.33	1.17

Appendix

A number of highly detailed tables may be of interest to some readers and are thus given in this Appendix. They are as follows:

Table A.1: Data Availability and HUD-Adjusted Area Median Family Income (HAMFI), by Metropolitan Area.

Table A.2: American Housing Survey: Descriptive Statistics.

Table A.3: Comparison of Incomes between American Housing Survey and National Data by Regions, 1985 and 1993.

Table A.4: Incomes and Household Characteristics of Kernels in Moderately High Income Neighborhoods.

Table A.5: Probability of Living in Neighborhoods with Incomes above the Middle Level, by Type of Urban Area.

Table A.6: Incomes of Kernels and Income Mixing in Their Neighborhoods.

Table A.7: Neighborhood Income Distribution Measures, All Kernels and Neighborhoods.

Table A.8: Kernels in Neighborhoods with Low Incomes (Median less than 80% HAMFI).

Table A.1:
Data Availability and HUD-Adjusted Area Median Family Income (HAMFI), by Metropolitan Area

Metropolitan area names are abbreviated by retaining the first word of their standard names. HAMFI is as supplied by HUD. Data availability is reported on the basis of the authors' own processing using the AHS data from the National Core and Supplement CD-ROM [US Bureau of the Census (1996)]. Data for suppressed names are not reported.

Name of MSA	1985			1993	
	Clusters	Observations	HAMFI ($)	Clusters	Observations
Akron, OH	.	.	29902	1	10
Albany, NY	2	22	32661	3	28
Albuquerque, NM	2	20	28038	3	31
Allentown, PA	2	22	32065	2	22
Alton, IL	.	.	32587	.	.
Anaheim, CA	10	94	42131	14	127
Appleton, WI	.	.	31617	1	9
Atlanta, GA	11	88	34675	11	103
Atlantic City, NJ	.	.	32363	.	.
Augusta, GA-SC	.	.	27814	1	11
Austin, TX	1	9	31170	1	11
Bakersfield, CA	.	.	26099	.	.
Baltimore, MD	7	58	35197	10	101
Baton Rouge, LA	1	11	27889	2	21
Beaumont, TX	.	.	26547	.	.
Beaver, PA	1	8	24384	2	19
Bergen, NJ	2	18	43623	5	56
Birmingham, AL	4	41	25801	6	59
Boston, MA	8	78	38179	16	149
Boulder, CO	1	10	37657	1	11
Bridgeport, CT	3	30	37434	4	43
Canton, OH	1	10	27591	2	22
Charleston, SC	.	.	26397	.	.
Chattanooga, TN-GA	1	9	24981	2	16
Chicago, IL	21	196	35495	29	248
Cincinnati, OH-KY-IN	2	18	31617	3	34
Cleveland, OH	5	54	31543	9	90
Colorado Springs, CO	1	11	29156	1	6
Columbia, SC	.	.	29828	1	11
Columbus, OH	1	10	31319	2	17
Corpus Christi, TX	1	6	24459	2	11
Dallas, TX	6	56	33929	11	102
Davenport, IL	4	34	28933	4	35

Name of MSA	1985			1993	
	Clusters	Observations	HAMFI ($)	Clusters	Observations
Daytona Beach, FL	1	8	24682	1	19
Denver, CO	1	9	34600	1	15
Des Moines, IA	.	.	31841	.	.
Detroit, MI	17	172	34968	25	242
Duluth, MN-WI	1	5	26099	1	5
East St Louis, IL	.	.	32587	.	.
El Paso, TX	1	10	20656	4	38
Erie, PA	.	.	26621	.	.
Eugene, OR	.	.	25279	1	11
Evansville, IN-KY	.	.	28485	.	.
Flint, MI	1	11	31244	2	21
Fort Lauderdale, FL	2	17	30648	4	32
Fort Myers, FL	.	.	26919	.	.
Fort Wayne, IN	.	.	31692	.	.
Fort Worth, TX	4	37	32288	5	47
Fresno, CA	1	11	24608	2	12
Gary-Hammond, IN	1	11	31394	2	24
Grand Rapids, MI	2	20	32810	4	43
Greensboro, NC	.	.	29306	1	7
Greenville, SC	1	11	27665	1	11
Hartford, CT	.	.	36464	1	9
Honolulu, HI	1	10	37210	1	10
Houston, TX	10	84	31692	16	146
Indianapolis, IN	2	21	31916	2	15
Jackson, MS	.	.	25875	.	.
Jacksonville, FL	2	20	28560	5	34
Jersey City, NJ	3	32	29231	5	50
Johnson City, TN-VA	2	22	22296	2	24
Kansas City, MO-KS	4	43	32214	8	71
Knoxville, TN	3	30	24981	4	37
Lake County, IL	.	.	35495	.	.
Lakeland, FL	1	11	24160	1	8
Lancaster, PA	.	.	31319	.	.
Lansing, MI	1	10	32885	1	3
Las Vegas, NV	2	22	30200	2	22
Lawrence, MA-NH	.	.	37891	.	.
Lexington, KY	.	.	27665	1	10
Little Rock, AR	1	9	27143	2	11
Los Angeles, CA	33	317	32065	46	446
Madison, WI	.	.	35271	1	16
McAllen, TX	1	7	15138	1	5

Name of MSA	1985			1993	
	Clusters	Observations	HAMFI ($)	Clusters	Observations
Melbourne, FL	.	.	29529	1	1
Memphis, TN-AR-MS	3	28	26323	5	46
Miami, FL	7	60	25950	13	101
Middlesex,	2	21	45711	3	31
Milwaukee, WI	1	10	33109	4	41
Minneapolis, MN	8	82	36986	12	116
Mobile, AL	4	35	23116	4	38
Modesto, CA	.	.	27069	.	.
Monmouth, NJ	3	15	38627	4	26
Montogomery, AL	.	.	26248	1	8
Nashville, TN	1	12	29306	2	24
Nassau, NY	12	115	47053	12	125
New Haven, CT	2	22	37434	3	30
New Orleans, LA	2	15	25130	5	42
New York City, NY	30	260	32904	47	442
Newark, NJ	5	48	41759	9	92
Norfolk, VA	5	50	29907	6	60
Oakland, CA	9	82	39074	11	99
Oklahoma City, OK	4	39	27814	4	37
Omaha, NE-IA	1	11	30946	1	11
Orlando, FL	1	9	29828	5	48
Oxnard, CA	1	11	41162	2	19
Pensacola, FL	1	10	24981	1	10
Peoria, IL	1	10	30946	1	9
Philadelphia, PA	19	194	34749	24	237
Phoenix, AZ	5	53	31021	7	53
Pittsburgh, PA	5	30	27367	10	82
Providence, RI	1	10	30957	1	8
Raleigh, NC	2	19	34675	3	31
Riverside, CA	8	80	30648	9	89
Rochester, NY	3	29	33854	3	30
Rockford, IL	.	.	31766	.	.
Sacramento, CA	4	30	31766	5	44
St Louis, MO-IL	5	46	32587	7	62
Salem, MA	.	.	38179	.	.
Salinas, CA	.	.	29753	.	.
Salt Lake City, UT	3	26	30275	4	40
San Antonio, TX	4	32	25726	5	47
San Diego, CA	6	62	32736	9	80
San Francisco, CA	8	70	40491	16	138
San Jose, CA	5	50	44219	8	81

Name of MSA	1985			1993	
	Clusters	Observations	HAMFI ($)	Clusters	Observations
Santa Barbara, CA	1	11	33929	1	5
Santa Rosa, CA	.	.	34451	.	.
Sarasota, FL	.	.	29455	.	.
Scranton, PA	2	22	26025	2	19
Seattle, WA	3	28	35793	10	100
Shreveport, LA	3	32	24086	3	27
Spokane, WA	2	22	26099	2	22
Springfield, MA	1	11	29455	1	11
Stamford, CT	1	6	37434	1	10
Stockton, CA	.	.	28485	.	.
Syracuse, NY	2	18	30872	3	31
Tacoma, WA	2	19	28858	3	31
Tampa, FL	7	57	26025	7	57
Toledo, OH	3	27	30424	4	33
Trenton, NJ	.	.	40193	.	.
Tucson, AZ	4	37	26621	5	62
Tulsa, OK	1	9	27963	2	24
Utica-Rome, NY	1	11	26248	1	9
Vallejo, CA	1	10	34898	2	19
Washington, DC	14	142	45189	22	205
Waterbury, CT	.	.	37434	2	16
West Palm Beach, FL	4	28	32139	5	39
Wichita, KS	1	11	30722	3	32
Worcester, MA	2	19	38179	2	24
Youngstown, OH	2	20	26621	4	37

Table A.2:

American Housing Survey: Descriptive Statistics

Household data and dwelling units data are for units interviewed both in 1985 and 1993.

	Mean		Coefficient of Variation	
Cluster-Survey Data	1985	1993	1985	1993
Cluster-averaged data	662	984		
Household income ($)	29407	36932	.577	.558
CPI-Urban (all)	107.6	144.5		
Monthly rent ($)	348	490	.466	.455
Property value ($)	75986	104923	.616	.691
CPI-Urban (housing)	107.7	141.2		
Household data	7322	7322		
Date head moved in (19 - -)	74.9	81.5	.155	.153
Age of head (years)	48.46	49.64	.363	.355
Highest grade (years)	12.54	12.92	.441	.410
Race (% White)	71	67		
Household size	2.62	2.56	.574	.580
Household income ($)	29433	37436	.846	.844
Dwelling unit data	7322	7322		
Number of rooms	5.35	5.39	.353	.341
Unit area (ft^2)	1565.1	1559.7	.603	.564
Appreciation rate$_{t,t-1}$ (owners)	.0601	.0264		
Monthly rent (renters)	326.5	465.1	.519	.483
Property value ($, owners)	79800	110303	.679	.727
Publicly owned housing (PROJ=1)	3.1%	2.6%		
National sample	43104	49326	43104	49326
Household income	28330	37693	.849	.849
Owner-occupant (%)	63.69	63.81	.755	.753
Age of head (years)	48.15	49.01	.365	.356
Highest grade (years)	12.38	12.87	.271	.244
Race (% White)	86.7	84.67	.391	.425

Table A.3:

Comparison of Incomes between American Housing Survey and National Data by Regions, 1985 and 1993

Sample: Kernels and Neighbors

U.S.- designated statistics are obtained from the *Statistical Abstract of the United States* [U.S. Bureau of the Census (1987; 1995)] and apply to the entire U.S. and regions, as appropriate, and not just urban areas. U.S. median housing costs and property values also apply to the entire U.S. and regions and are obtained from the AHS [U.S. Bureau of the Census (1985; 1993)]. All other statistics are based on authors' own processing of the American Housing Survey data [U.S. Bureau of the Census (1996)].

Year	1985					1993				
Regions	All SMSAs	Mid West	North East	South	West	All SMSAs	Mid West	North East	South	West
Summary Statistics										
Mean Income ($)	29410	26658	31140	28934	30928	37490	34085	41001	35893	39470
CV Income	.846	.818	.858	.859	.827	.854	.849	.850	.874	.819
Median Income ($)	23000	21700	24000	22145	24565	28248	26000	30075	26312	30336
U.S. Mean Income ($)	29066	28149	31146	27044	31475	41428	39442	45319	38249	45284
U.S. Median Income ($)	23618	23551	25485	21397	25782	31241	31400	33747	28441	33739
Mean HAMFI ($)	30845	31001	34267	27608	32381	41365	41574	45954	37023	43425
U.S. Median Monthly Housing Costs ($)	357	330	388	322	427	487	424	551	445	579
U.S. Median Property Values ($)	63211	45108	76224	47310	81913	86529	71898	116102	70376	134430

Table A.4:
Incomes and Household Characteristics of Kernels in Moderately High Income Neighborhoods

Table reports the probability that a kernel household with selected household and/or dwelling structure characteristics will live in a moderately high income neighborhood (median neighborhood income greater than 120% of HUD-Adjusted Median Family Income for the metropolitan area).

In the left margin, relative income categories are defined in terms of kernel household income as percentage of HUD-Adjusted Area Median Family Income. Columns 1 and 7 report percentiles for each of those income categories for 1985 and 1993, respectively. Columns 2–6 and Columns 8–12 give the probabilities that a kernel household with specified characteristics and with income in the respective category lives in a neighborhood with specified characteristics. Specifically, The first row of Column 2 gives the probability that a kernel household with extremely low income (less than 30% of HAMFI) who is a subsidized renter lives in a neighborhood where at least the richest one-half of residents have incomes above 120% of HAMFI; successive rows report the probability of such households with very low income (30%–50% of HAMFI), low income, etc. Column 3 gives the probability that a kernel household who is unsubsidized renter and has extremely low income (less than 30% of HAMFI), very low income (30%–50% of HAMFI), etc., lives in a neighborhood where at least the richest one-half of residents have incomes above 120% of HAMFI. Column 4 gives the probability that a kernel household who is an unsubsidized owner with extremely low income (less than 30% of HAMFI), with very low income ((30%–50%) of HAMFI), etc., lives in a neighborhood where at least the richest one-half of residents have incomes above 120% of HAMFI. Column 5 gives the probability that a kernel household lives in a single-unit structure with extremely low income (less than 30% of HAMFI), with very low income ((30%–50%) of HAMFI), etc., and in a neighborhood where at least the richest one-half of residents have incomes above 120% of HAMFI. Column 6 gives the probability that a kernel household who lives in a structure with at least 20 units and has extremely low income (less than 30% of HAMFI), very low income (30%–50% of HAMFI), etc., lives in a neighborhood where at least the richest one-half of residents have incomes above 120% of HAMFI.

Year	1985						1993					
	All	Kernel			Structure		All	Kernel			Structure	
Relative Income		Sub/d Renter	Unsub/d Renter	Owner	Single Fam.	> 20 Units		Sub/d Renter	Unsub/d Renter	Owner	Single Fam.	> 20 Units
% of HAMFI	1	2	3	4	5	6	7	8	9	10	11	12
0–30%	16	0	7	1	8	0	17	7	3	15	11	4
30%–50%	13	0	6	11	10	0	16	0	0	11	9	0.0
50%–80%	16	0	4	24	22	0	18	0	3	13	13	9
80%–100%	13	0	4	21	21	0	12	0	9	20	20	0.0
100%–120%	9	0	0	18	19	0	9	0	7	14	12	33
120%–150%	10	0	0	30	29	0	10	0	11	20	23	14
150%–200%	11	0	22	33	41	50	10	0	18	32	32	26
200%–	12	0	18	60	57	68	11	0	38	50	51	33

Table A.5:

Probability of Living in Neighborhoods with Incomes above the Middle Level, by Type of Urban Area

For kernel households with incomes in each of the categories of the first column who live in a particular type of urban area, the entries in the table give the probability that such a kernel household lives in a neighborhood where one-half of incomes are above the middle level (median neighborhood income greater than 80% of HAMFI).

Year	1985			1993		
Kernel Income as % of HAMFI	Central City	Urban Nonmetro	Suburban	Central City	Urban Nonmetro	Suburban
0–30%	14	0	25	15	0	28
30%–50%	12	29	22	22	22	36
50%–80%	22	54	46	25	13	32
80%–100%	41	60	68	26	50	67
100%–120%	47	60	65	53	64	60
120%–150%	62	57	80	75	55	70
150%–200%	74	71	79	66	70	73
200%–	64	67	86	71	56	78

Table A.6:

Incomes of Kernels and Income Mixing in Their Neighborhoods

This table gives the probability that a kernel household lives in a neighborhood with median income belonging to respective HAMFI-based quantiles. For example, the probability that a kernel household with income less than 30% of HAMFI lives in a neighborhood whose median income is also less than 30% of HAMFI is 42.7% in 1985.

Sample: all kernels and their neighborhoods

Income of Kernel	1985 Neighborhood Median Income as % of HAMFI							
as % of HAMFI	0-30%	30-50%	50-80%	80-100%	100-120%	120-150%	150-200%	200%-
0–30%	42.7	22.5	14.4	3.0	7.0	8.3	0.0	0.0
30%–50%	20.0	26.8	17.0	5.9	1.8	6.7	3.2	4.0
50%–80%	17.3	26.8	20.3	13.4	14.0	18.3	3.2	8.0
80%–100%	5.3	7.0	13.7	20.8	19.3	11.7	12.9	4.0
100%–120%	6.7	7.0	11.1	13.4	17.5	10.0	9.7	4.0
120%–150%	0.0	4.2	9.8	16.8	15.8	15.0	12.9	4.0
150%–200%	5.3	2.8	5.9	17.8	15.8	15.0	32.3	12.0
200%–	2.7	2.8	7.8	7.9	8.8	15.0	25.8	64.0
Total	100	100	100	100	100	100	100	100

Income of Kernel	1993 Neighborhood Median Income as % of HAMFI							
as % of HAMFI	0-30%	30-50%	50-80%	80-100%	100-120%	120-150%	150-200%	200%-
0–30%	50.5	25.4	14.5	7.0	7.0	9.9	5.3	5.0
30%–50%	17.5	22.3	12.2	12.5	6.0	4.2	5.3	5.0
50%–80%	17.5	20.8	26.6	12.5	8.0	11.3	7.9	10.0
80%–100%	6.8	15.4	11.7	18.0	11.0	15.5	13.2	0.0
100%–120%	2.9	10.0	7.9	16.4	15.0	9.9	5.3	0.0
120%–150%	1.0	1.5	9.8	17.2	19.0	9.9	13.2	10.0
150%–200%	1.9	1.5	9.8	10.9	20.0	22.5	15.8	10.0
200%–	12.8	3.1	7.5	5.5	14.0	16.9	34.2	60.0
Total	100	100	100	100	100	100	100	100

Table A.7:

Neighborhood Income Distribution Measures, All Kernels and Neighborhoods

This table shows the distribution of specified measures for neighborhoods according to kernel household income relative to HAMFI. The specified measures are: the minimum income of all neighbors, the 25%-tile (meaning, the value which defines the bottom 25% of incomes within the neighborhood, referred to in the table as Q_{25}), the median, the 75%-tile (meaning, the value which defines the top 25% of incomes within the neighborhood, referred to in the table as Q_{75}), and the maximum income of all neighbors. For example, column 1 shows that the minimum income in a neighborhood will fall within 0-30% of HAMFI with probability 66.1% in 1985, it will fall within 30-50% with probability 18.0%, and so on.

Year	1985					1993				
Neighborhood Measure	Neighborhood Distribution Statistic					Neighborhood Distribution Statistic				
as % of HAMFI	Min	Q_{25}	Median	Q_{75}	Max	Min	Q_{25}	Median	Q_{75}	Max
0–30%	66.1	31.1	13.1	5.1	1.4	68.5	31.5	12.8	5.2	2.4
30%–50%	18.0	21.6	12.4	7.0	3.5	19.5	27.4	16.2	6.7	2.1
50%–80%	11.9	25.8	26.7	16.6	6.1	8.2	24.0	26.6	18.0	8.3
80%–100%	2.3	10.5	17.6	13.4	8.4	1.7	8.5	15.9	12.1	6.1
100%–120%	0.0	3.8	9.9	15.5	8.2	1.1	4.4	12.4	14.9	9.6
120%–150%	1.6	4.0	10.5	14.8	14.0	0.9	2.6	8.8	18.9	14.9
150%–200%	0.0	3.0	5.4	16.6	22.0	0.0	1.4	4.7	15.1	24.1
200%–	0.0	0.0	4.4	11.0	36.5	0.0	0.4	2.5	9.1	32.5
Total	100	100	100	100	100	100	100	100	100	100

Table A.8:

Kernels in Neighborhoods with Low Incomes (Median Less than 80% HAMFI)

Table reports the probability that a kernel household with selected household and/or dwelling structure characteristics will live in a neighborhood with extremely low and low incomes (median neighborhood income less than 80% of HUD-Adjusted Median Family Income for the metropolitan area).

In the left margin, relative income categories are defined in terms of kernel household income as percentage of HUD-Adjusted Area Median Family Income. Column 1 reports percentiles for each of those income categories in the entire population of kernels, for 1985 and 1993, respectively. Column 2 reports the percentage of kernels in each income category who live in neighborhoods where the poorest one-half of residents have incomes less than 80% of HAMFI. Columns 3–5 give the probabilities that a kernel household with specified characteristics and with income in the respective category lives in a neighborhood with specified characteristics. Specifically, the first row of Column 3 gives the probability that a kernel household with extremely low income (less than 30% of HAMFI) who is a subsidized renter lives in a neighborhood where the poorest one-half of residents have incomes less than 80% of HAMFI; successive rows report the probability of such households with very low income (30%–50% of HAMFI), low income, etc. Column 4 gives the probability that a kernel household who is unsubsidized renter and has extremely low income (less than 30% of HAMFI), very low income (30%–50% of HAMFI), etc., lives in a neighborhood where at least the richest one-half of residents have incomes above 120% of HAMFI. Column 5 gives the probability that a kernel household who is an owner with extremely low income (less than 30% of HAMFI), with very low income ((30%–50%) of HAMFI), etc., lives in a neighborhood where the poorest one-half of residents have incomes less than 80% of HAMFI. Column 6 gives the probability that a kernel household lives in a single-unit structure with extremely low income (less than 30% of HAMFI), with very low income ((30%–50%) of HAMFI), etc., and in a neighborhood where the poorest one-half of residents have incomes less than 80% of HAMFI. Column 7 gives the probability that a kernel household who lives in a structure with least 2 but less than 19 units and has extremely low income (less than 30% of HAMFI), very low income (30%–50% of HAMFI), etc., lives in a neighborhood where the poorest one-half of residents have incomes less than 80% of HAMFI. Column 8 gives the probability that a kernel household who lives in a structure with least 20 units and has extremely low income (less than 30% of HAMFI), very low income (30%–50% of HAMFI), etc., lives in a neighborhood where the poorest one-half of residents have incomes less than 80% of HAMFI.

Year	1985							
		% Kernels in Neighborhoods with Median <80% HAMFI						
	Population	Kernels	Kernel Tenure			Kernel Structure		
Relative Income	Distribution of Kernels	All	Sub/d Renter	Unsub/d Renter	Owner	Single Fam.	2–19 Units	20+ Units
% of HAMFI	1	2	3	4	5	6	7	8
0–30%	14.3	85.4	100.0	80.0	77.7	80.5	85.7	100.0
30%–50%	12.7	86.2	100.0	85.7	77.8	77.5	85.2	100.0
50%–80%	17.3	63.6	100.0	82.2	35.1	46.6	90.3	87.5
80%–100%	12.9	39.6		69.2	25.0	32.8	71.4	50.0
100%–120%	10.6	44.3	100.0	81.9	34.7	36.2	55.6	100.0
120%–150%	10.1	31.0		57.1	22.7	25.0	57.1	66.7*
150%–200%	11.2	23.4		38.9	17.4	18.4	38.5	50.0*
200%–	10.8	25.8		63.7	13.3	16.7	71.4	23.1
kernels in category as % of all	100	52.2	100.0	73.7	36.4	39.6	76.9	82.0
kernel frequency in category	100	100	5.2	33.2	61.5	76.4	25.0	7.7
Number of kernels	573	299	30	190	352	386	143	44

* Observations ≤ 3.

WORKING PAPER SERIES 2004

ase.tufts.edu/econ/papers

2004-01 EGGLESTON, Karen & Winnie Yipp; Hospital Competition under Regulated Prices: Application to Urban Health Sector Reforms in China

2004-02 EGGLESTON, Karen & Chee-Ruey Hsieh; Health Care Payment Incentives: A Comparative Analysis of Reforms in Taiwan, Korea and China

2004-03 EGGLESTON, Karen; Competition, Altruism, and Provider Payment

2004-04 LOURY, Linda; Some Job Contacts are More Equal Than Others: Earnings and Job Information Networks

2004-05 LOURY, Linda; Teen Childbearing and Community Religiosity

2004-06 LOURY, Linda; Religiosity and High School Misbehavior

2004-07 MUKAND, Sharun & Sumon Majumdar; Policy Gambles

2004-08 MUKAND, Sharun & Sanjay Jain; Public Opinion and the Dynamics of Reform

2004-09 MUKAND, Sharun, Sumon Majumdar & Anandi Mani; Politics, Information and the Urban Bias

2004-10 BIANCONI, Marcelo; Aggregate and Idiosyncratic Risk and the Behavior of Individual Preferences under Moral Hazard

2004-11 BIANCONI, Marcelo; Transfer Programs and Consumption under Alternative Insurance Schemes and Liquidity Constraints

2004-12 BIANCONI, Marcelo; Heterogeneity, Adverse Selection and Valuation with Endogenous Labor Supply

2004-13 BIANCONI, Marcelo; The Welfare Gains from Stabilization in a Stochastically Growing Economy with Idiosyncratic Shocks and Flexible Labor Supply

2004-14 SHIMSHACK, Jay P. and Michael B. Ward; Enforcement and Environmental Compliance: A Statistical Analysis of the Pulp and Paper Industry

2004-15 NORMAN, George and Joel P. Trachtman; The Customary International Law Supergame: Order and Law

2004-16 BROWN, Drusilla K. and George Norman; Optimal Intellectual Property Rights Protection and Humanitarian Assistance during and International Health Emergency

2004-17 LOURY, Linda; Job Tenure and Personal Contacts: Good Matches or Limited Choices?

2004-18 KUTSOATI, Edward and Sharun Mukand; Expectations and the Central Banker: Making Decisions the Market Expects to See?

2004-19 RICHARDS, Daniel; Price Discrimination and the Long Boom.

2004-20 HARDMAN, Anna and Yannis Ioannides; Income Mixing and Housing in U.S. Cities: Evidence From Neighborhood Clusters of the American Housing Survey.

2004-21 HARDMAN, Anna and Yannis Ioannides; Neighbors' Income Distribution: Economic Segregation and Mixing in US Urban Neighborhoods. [revised]

2004-22 IOANNIDES Yannis and Andrew Eggers; The Role of Output Composition in the Stabilization of the U.S. Output Growth

Year	1993							
		% Kernels in Neighborhoods with Median <80% HAMFI						
	Population	Kernels	Kernel Tenure			Kernel Structure		
Relative Income	Distribution of Kernels	All	Sub/d Renter	Unsub/d Renter	Owner	Single Fam.	2–19 Units	20+ Units
% of HAMFI	1	2	3	4	5	6	7	8
0–30%	17.7	81.6	96.9	87.1	63.0	75.4	90.7	78.3
30%–50%	12.6	72.2	84.3	82.0	64.4	63.2	87.5	100.0
50%–80%	17.3	73.4	100*	81.1	66.6	66.0	89.2	81.9
80%–100%	12.7	51.0	100*	65.8	40.7	39.3	67.9	100
100%–120%	9.7	42.4		53.5	34.0	37.9	58.9	33.3*
120%–150%	9.8	30.4	100.0*	57.9	20.4	22.8	46.7	57.2
150%–200%	10.3	28.1	100.0*	58.8	21.5	24.6	66.7	50.0
200%–	99.9	27.5	100.0*	25.0	28.2	25.6	60.0*	33.3*
kernels in category as % of all	100	55.6	95.3	71.0	42.6	45.5	78.3	72.2
kernel frequency in category	100	100	5.4	34.3	60.3	67.5	23.5	9.0
Number of kernels	804	447	43	274	482	543	189	72

* Observations ≤ 3.