Changes in racial and ethnic diversity in neighborhoods in large urban areas in the U.S., 1980-2020

John R. Ottensmann
Indiana University Indianapolis
john.ottensmann@gmail.com
urbanpatternsblog.wordpress.com
March 2024

Abstract

The shares of the population of census tracts white, black, Latino, and Asian are used to calculate a measure of neighborhood racial and ethnic diversity for the tracts in 56 large urban areas from 1980 to 2020. Diversity has increased dramatically over the period, with about three-quarters of all tracts increasing in each decade. Multilevel models are used to evaluate the association of tract and urban area characteristics with the change in tract diversity from 1980 to 1990 and from 2010 to 2020. Tract characteristics account for a much larger share of the variation in change in diversity than urban area characteristics. The change in the tract percent white is negatively related to diversity change while the percent white is positively related. For the other three groups the directions of the relationships are reversed. Tract diversity at the start of the period is negatively related to change in diversity (more opportunity for increase?) while urban area diversity is positively related (greater tolerance for diversity?). Population growth is positively and significantly related to diversity change, but the effect is smaller.

Introduction

The study of racial and ethnic change in America's neighborhoods has been shaped (and limited) by a focus on blacks and whites, on large urban areas in the northeast and midwest, and on change in the 1940s, 1950s, and 1960s, an argument also made by Wood and Lee (1991). This has led to the idea that racial change involves blacks moving into all-white neighborhoods which inevitably leads to those areas becoming all black. Increasing evidence suggests that this is not always the case, and that conditions specific to those urban areas and that period contributed to that form of racial change.

The increasing racial and ethnic diversity of large cities has gradually led to the realization that the exclusive focus on whites and blacks was insufficient to capture the full scope of racial and ethnic diversity. Most of the studies of neighborhood racial and

ethnic change for multiple groups classify census tracts into a number of groups based on their racial and ethnic composition. The earlier studies developed these classes using the percentages of tract populations in each of the groups, producing significant numbers of classes reflecting various combinations. Some later studies used indexes of diversity to classify tracts, producing fewer groups. In both cases change is assessed by the comparison of tract classes at the start and end of a period, most often the decade between censuses. Such classification is necessarily arbitrary and involves the loss of information from the underlying measures. For example, a tract can experience a large change in percent black that keeps the tract in the same class and therefore is not taken as evidence of change while another tract can have a small change near a class boundary that moves the tract to another class and is seen as evidence of change.

This research examines neighborhood racial and ethnic change in 56 large urban areas from 1980 to 2020. Census tract populations in the largest racial and ethnic groups—whites, blacks, Latinos, and Asians—are used. An index of diversity summarizes the composition of the population for each tract. Rather than classifying tracts and examining class stability and change over time, the amount of change in the diversity index is used directly as the measure of change.

The next section provides a brief review of the literature on racial and ethnic change in cities in the United States. The following section outlines the basic dataset being used to define the urban areas that serve as the context for the research. The racial and ethnic groups and data are described next, followed by the presentation of the index employed as the measure of neighborhood diversity. Results are provided in two sections. The first gives a basic description of racial diversity and change in the census tracts and the urban areas over time. Multilevel models are used to examine how characteristics of both the census tracts and the urban areas are associated with changes in the index of neighborhood diversity.

The evolution of the the study of racial and ethnic neighborhood change

Early discussions viewed racial change as an example of a more general ecological process of invasion and succession that explains neighborhood change (Burgess 1928; Park 1936). One of the first detailed studies of racial change at the census tract level by Duncan and Duncan (1957) looks at Chicago from 1940 to 1950 using this framework. They find that virtually all of the tracts with some black occupancy in 1940 had become nearly all black by 1950.

In an important, comprehensive study of black segregation and neighborhood change, Taeuber and Taeuber (1965) likewise employ the concept of invasion and succession, though they found that succession did not take place in all census tracts. Aldrich (1975) reviews the literature on invasion and succession and is skeptical that mixed neighborhoods would inevitably experience complete succession. He makes the

point that much research had focused on change in the 1950s and 1960s in cities experiencing large black inmigration from the South. Massey (1983) continues to use the framework in the analysis of residential change involving the Hispanic population.

An alternative view of neighborhood change from white to black is the "tipping-point hypothesis" described by Grodzins (1957, 1958). This posits that when the percentage black in a neighborhood exceeds some critical threshold, whites would move out and the area would become all black. Wolf (1963) adds the observation that such "white flight" is not required. Households regularly move for reasons other than race. All that is required for a neighborhood to "tip" is that white households refuse to move into a neighborhood once the percentage black exceeds the critical level.

Schelling (1971) provides the most sophisticated analysis of the phenomenon, arguing that white tolerance for blacks could vary widely and still result in neighborhood instability and tipping. As the percentage black in a neighborhood increases, the least tolerant whites would leave, causing the neighborhood to become more black. This would result in the next least tolerant whites leaving, and so on until the neighborhood becomes largely black.

Studies proceeded to search for evidence to support or refute the tipping point hypothesis. Goering (1978) reviews this literature and concludes that change from white to black was often a continuous process. Neighborhoods vary and one could not conclude that there is always a tipping point, though there may be such a critical level in some areas. Research has continued to yield conflicting results. For example, two papers by economists (Card, Man, and Rothstein 2008; Easterly 2009) use different sophisticated econometric methods *applied to the same dataset* and come to opposite conclusions.

Numbers of studies have countered the assumption of succession and tipping that neighborhoods with some blacks would inevitable become all black. Taeuber and Taeuber (1965) look at racial change in 10 cities, essentially replicating the Duncan and Duncan (1957) study of Chicago. They find few or no stable tracts in the cities they studied in the Northeast and Midwest but significant numbers of such tracts in cities in the South.They conclude that

> Although an aura of inevitability is often attributed to this process [of residential succession], as in the popular notion of a "tipping point," processes of change in the racial composition of neighborhoods do not always follow such simple patterns.

(Was it the case that their early findings had limited effect on subsequent consideration of racial change because what was happening in cities in the South was deemed to be less important?)

Lee and Wood (1991) conclude racial succession was far from inevitable in racially-mixed tracts in 58 central cities from 1970 to 1980. They do find large variations by region, however, with succession most common in Midwestern cities and stability most prevalent in cities in the West. Another study (Wood and Lee 1991) examines racial change in five large cities from 1940 to 1980. In four of the five cities, racial succession was lower in the later decades while stability increased. New Orleans was the exception.

The existence of somewhat greater neighborhood racial stability in later decades suggests that the large migration of blacks to northern cities through the 1960s could be a possible explanation for the conclusions drawn earlier that racial succession was virtually inevitable. In a study of the housing market in four racially mixed neighborhoods in Philadelphia, Rapkin and Grigsby (1960) find significant numbers of home purchasers to be white. They suggest that one factor causing racial transition was the large numbers of blacks moving into the city, creating a high demand for housing by black residents. Taeuber and Taeuber (1965) cite "the greater the rate of Negro population growth relative to white population growth" as a cause of racial succession. Goering (1978) likewise mentions this in his discussion of the literature on the tipping point. A colleague and I (Ottensmann and Gleeson 1992) examine the race of households that recently move into census tracts in the Chicago metropolitan area in 1960, 1970, and 1980. The percent of households moving into racially mixed neighborhoods that were white is surprisingly high and increased over time for neighborhoods with all but the highest percent black.

The findings that racial succession is not always inevitable has produced studies of the presence of stable racial integration. This begins with case studies of specific neighborhoods that have remained integrated over a period of time. Taub, Taylor, and Dunham (1984) and Maly (2005) are significant book-length efforts, each examining multiple neighborhoods. The most comprehensive study of neighborhood racial change and stable racial integration is Ellen (2000). She examines change at the census tract level in 34 large metropolitan areas from 1970 to 1990 and considers various aspects of that change in great detail.

The very large increases in racial and ethnic diversity in American cities in recent decades has made the exclusive focus on racial change involving only blacks and white increasingly untenable. Of course some cities have long had large populations of other groups. From early on, studies have sidestepped this issue by excluding cities with significant populations of groups other than blacks and whites (e.g., Taeuber and Taeuber 1965). But this practice has been continued more recently by those choosing to focus exclusively on black and white racial change, for example Ellen (2000) and Rawlings, *et al.* (2004).

The evolution of widening the focus to include more diverse populations can be seen in a series of three articles addressing stable racial integration that appear together

in a journal issue. Each establishes criteria for considering a neighborhood to be integrated. Smith (1998) considers only the percentages of the population black and white, explicitly excluding Hispanics. Galster (1998) acknowledges the importance of other racial and ethnic groups and proposes to deal with this by using percentages white and nonwhite, with the latter including members of other groups in addition to blacks. Finally, Ellen (1998) includes non-Hispanic whites, blacks, and members of all other racial and ethnic groups taken together.

The typical approach for looking at racial transition involving blacks and whites has been to use percentages to classify tracts as largely black, largely white, and between these extremes, possibly into multiple groups. Racial change is then taken to be the transition from one group to another over an interval of time. A comparable approach uses the percentages in more racial and ethnic groups to classify tracts and examining transitions. But the use of more than two groups greatly adds to the complexity of the analysis.

An early example is the Denton and Massey (1991) study of change from 1970 to 1980 in census tracts in 60 metropolitan areas. They start with numbers of non-Hispanic whites, blacks, Hispanics, and Asians and classify tracts as all white or, if not, by the presence (30 or more persons) of members of the other groups. Separate classes are identified for the presence of one group only and for all possible combinations of two or three groups, giving eight classes. This produces an eight-by-eight transition matrix of racial and ethnic change over the decade. To take into account the relative presence of nonwhites, tracts are classified into eight additional classes using percentage ranges. Transitions among these classes are then presented for the tracts in each of the original classes.

Alba, *et al.* (1995) focuses on racial change in the New York region from 1970 to 1990, classifying tracts using percentages in the same four groups. Tracts in which a single group is dominant create four classes. Seven more classes have two, three, or four groups considered dominant, giving a total of 13 classes and a large transition matrix. Logan and Zhang (2010) study racial and ethnic change in 24 diverse metropolitan regions from 1980 to 2000. They classify tracts in a similar manner but include all possible combinations for a set of 15 classes and an even larger transition matrix.

As it became imperative to consider racial and ethnic groups beyond blacks and whites, similar problems of complexity emerged in the study of residential segregation. The original measures of segregation compare the residential distributions of two groups. Early efforts to examine segregation among a wider set of groups proceeded to compute these measures for all pairs of racial and ethnic groups. Massey and Denton (1987) produce 16 measures of segregation for four racial and ethnic groups for each of two years. An alternative approach uses measures of multi-group segregation that produce a single measure from information on the distribution of more than two groups

across census tracts. Examples include Iceland (2004), Farrell (2008), and Brown and Sharma (2010).

Similar approaches have been taken to examine the racial and ethnic change of neighborhoods over time. Maly (2000) develops a neighborhood diversity index and uses it to study change in the Chicago metropolitan area from 1980 to 1990. Tracts are classified using the values of the diversity index as integrated, moderately integrated, or segregated. The transition matrix shows the majority of tracts remaining in each class with small numbers having increased levels of diversity. Since this diversity index measures tract diversity in relation to diversity in the Chicago area, it is less useful for making comparisons across areas and over time.

Farrell (2005) and Farrell and Lee (2011) use a standard measure of diversity, the entropy index, for tracts in the 100 largest metropolitan areas in 1990 and 2000. The index ranges from 0 to 100, and the tracts are classified into five groups of equal width. Again, the largest numbers of tracts in each class in 1990, close to or over a majority, remain in the same class in 2000. Other tracts transition in either direction, to higher or lower diversity.

The studies focusing on racial change involving only blacks and whites classify neighborhoods as largely white or black or integrated to various degrees using the percentages in the tracts (e.g., Taeuber and Taeuber 1965; Wood and Lee 1991; Ellen 2000). Denton and Massey (1991), Friedman (2008), and Logan and Zhang (2010) address racial and ethnic change among more population groups, classifying neighborhoods into larger numbers of classes using the number or percentage in each of the groups. The alternative approach uses measures of neighborhood diversity to divide the neighborhoods into classes, as in Maly (2000), Farrell (2005), and Farrell and Lee (2011).

What all of these works have in common is that they use one or more continuous measures—counts or percentages in groups or an index of diversity—to classify census tracts into a limited number of classes. Such an approach has a number of problems. First, such classification involves a tremendous loss of the information contained in the original measures. Percentages range from 0 to 100. Even if rounded to whole numbers, they would represent 101 possible values compared to the relatively small number of classes resulting from these procedures.

Such classification involves the arbitrary specification of the classes. For example, in the case of black and white racial change, is a neighborhood to be considered integrated if the black presence ranges from 10 to 50 percent or from 10 to 90 percent. Both have been used. With studies including multiple groups, is a nonwhite group deemed to be present in a neighborhood if more than 30 are present? Or 100? Or if members of the group constitute 25 percent or more of the population? I am deliberately not providing citations because the purpose is not to critique any specific study but rather to highlight the arbitrary nature of the process.

Unless classes are very narrow and numerous, classification can produce results that are not especially reasonable. Consider a study of black-white change and stable racial integration in which tracts are classified as integrated if they include between 10 and 50 percent blacks. A neighborhood is considered to be stably racially integrated if from one time period to the next it remains in this class. If it moves to another class, say to greater than 50 percent black, it is not considered to be a stable tract. But this means that a tract changing from 49 to 51 percent black is taken to be not stable, but a tract changing from 11 to 49 percent black is an example of stable racial integration.

The classification procedure creates additional issues when tract populations or percentages in multiple groups are used to classify tracts. Even a rather coarse classification for each group can create a large number of classes when the combinations among the various groups are used to form the final classification. This results in large transition matrices showing the movement of tracts among the classes from one period to the next. Thus provides rich detail on the process of racial and ethnic neighborhood change. But at the same time this makes it difficult to draw more general conclusions about that change.

The studies using measures of multi-group diversity overcome this but they still rely on the classification of tracts into a smaller number of classes. When classes of equal width are specified, an argument can be made that the classification is somewhat less arbitrary, though the number of classes is of course arbitrary as is the decision to use equal width classes. One might just as logically divide the groups into classes using the quartiles of the diversity index.

Additional contributions to the study of these issues are surveys of individual preferences and willingness to move into neighborhoods with varying racial composition. An early important study of residents of the Detroit area (Farley, *et al.* 1978) presented respondents with diagrams illustrating neighborhoods with varying shares black and white and asked about their preferences. They conclude from the results that a tipping point of 30 percent black exists. The survey was repeated in four metropolitan areas (Farley, Fielding, and Krysan 1997) with findings of varying tolerance across the areas. Detroit was the least tolerant but still more tolerant than in the earlier survey. Multiple surveys have been conducted in southern California with varying conclusions drawn from the results (Clark 1992; Bobo and Zubrinsky 1996; Charles 2000).

The *Urban patterns 2* data

The *Urban patterns 2* dataset provides the context for this research. It includes housing unit counts for census tracts from 1950 to 2020 which are used to delineate 56 large urban areas in the United States for each census year. Data for 2010 and 2020 are from the census and the National Historical Geographic Information System (Manson,

et al. 2022). Data from the censuses from 1970 to 2000 are from a unique dataset from the Urban Institute and Geolytics (2003) with the data normalized to 2000 census tract boundaries. Housing units for 1950 and 1960 are estimated from the data on housing units by year built from later years, taking the numbers built before 1950 and 1960 as the estimates of the numbers present in those years. These estimates include error resulting from changes to the housing stock over time, especially the loss of units, but analyses suggest that the estimates for urban area totals are reasonable for two decades back in time. Census tract boundaries for 2020 are used for the dataset. The census tract relationship files are used to estimate values for the 2020 tracts from data for earlier years. Detailed documentation of the dataset and a listing of all data sources are provided in Ottensmann (2023a).

Urban areas consist of contiguous census tracts that meet urban criteria. Some large areas of continuous urban tracts include what should reasonably be considered two or more urban areas. Areas in the northeastern United States are a major example. To distinguish separate urban areas, Combined Statistical Areas (CSAs) are used (and Metropolitan Statistical Areas (MSAs) that are not included in a CSA). CSAs are used rather than the more commonly used MSAs as they better represent the full extent of urban areas. The CSAs are only used to identify the urban areas, such as Philadelphia, New York, and Hartford. The boundaries are established at the locations where the urban areas have become contiguous as they have expanded. The urban areas included in the dataset are the 56 areas with more than 300,000 housing units in 2020.

The criteria defining the urban areas are as close as possible to those being used for delineating the 2020 census Urban Areas, which include what were formerly called Urbanized Areas (U.S. Census Bureau 2022). A census tract is considered to be urban and is included in an urban area if it has a housing unit density greater than 200 housing units per square mile and is contiguous to the urban area. To include urban territory that is nonresidential, a tract is also included if over one-third of its area has impervious surface of 20 percent or more. An additional condition is that a tract is only considered to be urban if it has been designated as urban for the following census year, providing a pattern of cumulative expansion of the urban areas. This direction has been chosen rather than the reverse (if urban, then urban later) because the more recent data are considered to be more accurate.

Urban areas include multiple areas of urban territory that were originally separate but that have since growth together. Such areas that are sufficiently large are considered to be urban centers and are included in an urban area with tracts assigned to one of those urban centers. The Dallas-Fort Worth area is an example. As the areas become contiguous, tracts are assigned to the center growing more rapidly toward the other and to provide more continuous, less irregular boundaries. Areas are considered separate urban centers and are included in an urban area if the number of housing units in 2020 exceeds 16 percent of the total units in the urban area. This cutoff was

established by identifying as candidates all initially separate areas deemed large enough to potentially be considered urban centers and then setting the threshold. The smallest urban centers in relation to the total size of the urban area are Providence, with Boston; Tacoma, with Seattle; and High Point, with Greensboro and Winston-Salem. Next highest, at 11 percent are Port Charlotte in the Sarasota-Bradenton area and Winter Haven in the Orlando area. The names given to the urban areas include the names of all urban centers that are included.

Racial and ethnic groups and data

The identification of the racial and ethnic groups to be used in the research necessarily depends on the classifications used by the census for collecting and reporting the data. This section describes the selection of the four groups and the sources of the data.

The census considers Hispanic or Latino to be ethnic status, asking whether or not respondents identify themselves as members of that group. Those responding yes are all considered to constitute one of the racial and ethnic groups regardless of how they identify as to race. Those not Hispanic or Latino are then potential members of the other racial groups.

The three largest racial groups are those identifying only as whites, only as African Americans or blacks, and only as Asians. These are included in the research. The three other single-race groups (which include some other race) each have less than one percent of the United States population in 2020. Those specifying two or more races, allowed since the 2000 census, are only four percent of the population in 2020 and much smaller shares in the prior two censuses. As this was not an option for earlier censuses from which data are also used, this group could be included.

One modification is made for comparability with the earlier census data. Before 2000, the Native Hawaiian and Other Pacific Islander group was combined with the Asian group, which was identified as Asians and Pacific Islanders. So for the censuses from 2000 forward, the Asian group and the Native Hawaiian and Other Pacific Islander groups are combined to form an equivalent Asian and Pacific Islander group.

For this research, the distribution of the population by race and ethnicity is considered with respect to the total of the populations in the four groups. In other words, the (very small) population not in these groups is not included. Thus the percentages of the population in the four groups sums to one hundred percent.

Census tracts are taken to be the neighborhoods within which residential diversity is being considered. While tracts may not represent the extent of actual neighborhoods, they come closest among the census units for which the required data are available. They are the units most often used in studies of residential segregation and diversity. Also, the urban areas for each year have been delineated using census

tracts. Racial and ethnic group populations for tracts are thus also needed for aggregation to the larger areas.

Racial and ethnic group populations by census tract for 2000, 2010, and 2020 are from the census via the National Historical Geographic Information System (Manson, *et al.* 2022). The Neighborhood Change Database (Urban Institute and Geolytics 2003) is the source of the data for 1980 and 1990, normalized to 2000 census tract boundaries. The 1970 census was the first in which the question on Hispanic status was asked, but this was done in a way that produced data that was not reliable, and it is not used in this research (Cohn 2010). Hipp and Kim (2023) make the same decision.

For data using the 2000 and 2010 census tract boundaries (which change somewhat at each census), estimates are made for the 2020 tracts. Tract relationship files from the census are used for the estimation following the same procedure used for the urban patterns data as described in the detailed documentation (Ottensmann 2023a).

Several notes on presentation. The word "black" is used to refer to the group identifying as non-Hispanic African American or black. In a survey (Sigelman, Tuch, and Martin 2005) members of that group indicated approximately equal preference for the two terms. Black is chosen as the more long-standing descriptor. "Latino" is used for the Hispanic or Latino population. While more favor Hispanic over Latino, half indicate no preference (Pew Research Center 2013). Latino is more inclusive and accurate, encompassing those from Latin America whose native language is not Spanish. The group including Asians and native Hawaiians and other Pacific Islanders is labeled simply "Asians" for brevity and because they constitute the overwhelming majority in most cases. Reference to the more encompassing description will be made when appropriate (as in discussion of Honolulu). The groups are listed in this order, which is descending order of size at the start of the research in 1980.

Descriptions of areas as majority-minority or identification of groups other than whites as minorities will not be used in this paper. It becomes confusing and nonsensical to identify a group as a minority when in some situations they constitute a majority of the population. Furthermore, referring to whites as the majority group (when often they are not) and the other groups as minorities denotes a special status for whites as compared to the other groups, which is not appropriate. If the terms majority and minority are used, they will refer to the condition of constituting more than half or less than half of the population.

The measurement of neighborhood diversity and change

An index of diversity is presented that measures neighborhood diversity in census tracts based on the distribution of the population among the four racial and ethnic groups. This index is then used to produce the measure of change in racial and ethnic diversity. Three of the more useful reviews of measures of diversity are White

(1986), Reardon and Firebaugh (2002) and Budescue and Budescue (2012). All identify two diversity measures, the entropy index and the interaction index, called the generalized variance by Budescue and Budescue.

Entropy is a measure of uncertainty in information theory developed by Shannon (1948). This is related to diversity as the level of uncertainty about the group membership of an individual selected from a population is zero if the population consists only of members of a single group, minimum diversity. As diversity increases, so does the uncertainty. Theil (1972) applies entropy as a measure of diversity. The formula for the entropy index E is

$$E = \sum_{i=1}^{n} p_i \log\left(\frac{1}{p_i}\right) \tag{1}$$

where p_i is the proportion of the population in group i out of n groups and $\log(\,)$ is the natural logarithm. It tend to zero as the proportion in one group approaches 1 and increases to a maximum that depends on the number of groups when group proportions are equal.

The interaction index is the probability that two persons chosen at random from a population will be members of different groups. Simpson (1949) proposes the index as a measure of species diversity in an ecosystem. Herfindahl (1950) and Hirschman (1964) use the index as a measure of industrial concentration. The formula for the interaction index I is

$$I = 1 - \sum_{i=1}^{n} p_i^2 \tag{2}$$

The minimum value is zero when the entire population is concentrated in a single group. The maximum, which depends on the number of groups, is reached with equal proportions.

How these indexes measure diversity is not intuitive. A derivation of the interaction index is offered that is more transparent about how it serves as a measure of diversity. A criterion for a measure of diversity met by both the entropy and interaction indexes is that maximum diversity is achieved with equal proportions of the population in each of the groups (White 1986; Reardon and Firebaugh 2002; Budescu and Budescu 2012). This intuitive notion is taken as the starting point for developing the index of diversity. Equal proportions in all groups implies that the proportions are all equal to $1/n$, where n is the number of groups. Then the difference between the proportion in each group and $1/n$ is a measure of the contribution of that group to the departure from maximum diversity. These differences will be both negative and positive, so they are

squared and then summed to get a measure of the total departure from maximum diversity,

$$\sum_{i=1}^{n} \left(p_i - \frac{1}{n} \right)^2 \tag{3}$$

This sum can vary from zero for maximum diversity to 1 - 1/n when the entire population is concentrated in a single group. To normalize this to range from zero to one, multiply by the inverse of that maximum. Then subtract from one so the index increases with greater diversity, ranging from zero for concentration in a single group to one for equal proportions, maximum diversity. The formula for the diversity index D is then

$$D = 1 - \frac{n}{n-1} \sum_{i=1}^{n} \left(p_i - \frac{1}{n} \right)^2 = 1 - \frac{n}{n-1} \left(\sum_{i=1}^{n} {p_i}^2 - \frac{1}{n} \right) \tag{4}$$

The first expression shows the index as one minus the normed sum of the squared departures from maximum diversity. The second includes the sum of the squared proportions that is in the interaction index. And the term subtracted from one is that sum of squared proportions normalized to range from zero to one. So this diversity index is equivalent to the interaction index normed to vary from zero to one. This index can be used to calculate a measure of racial and ethnic diversity for any area including census tracts and the entire urban area.

The primary use of this index of diversity here is for determining the diversity of census tracts using the proportions of the population in each of the groups in the tract. The index is also used as a measure of the overall diversity of an urban area using the distribution among the groups in the urban area. For clarity of exposition, the diversity index values are multiplied by 100 to range from zero for no diversity, entire population in a single group, to 100, equal proportions in the four groups.

The use of an index of to calculate a single value for racial and ethnic diversity has been established in a number of papers in recent decades. But when used to look at change in neighborhood diversity, the index values have been used to classify census tracts into a small number of classes, with the problems discussed above. In this research, the measure of change is a single, continuous value, simply the value of the index of diversity at the end of a period minus the value at the start. Values greater than zero represent increases in diversity. Negative values are declines.

Neighborhood racial and ethnic diversity and change

This section sets the context for the research by first providing descriptions of the distribution of the population of the 56 urban areas among the racial and ethnic groups from 1980 to 2020. Summary statistics for the values of the diversity index for the census tracts and the changes for the tracts in these urban areas come next. Following is a brief look at the levels of tract diversity and change across the 56 urban areas.

Table 1 gives the distribution and change of the total population in the urban areas for the four racial and ethnic groups from 1980 to 2020. The top section presents the percentage in each group and the lower section gives the change in percent from year to year. The share of the population white declined dramatically, from 72 percent in 1980 to just under half of the population, 49.6 percent, in 2020. The decline was fairly steady, around five percentage points or slightly more in each decade. In contrast, the percent of the population black held steady at around 16 percent, with very little change in any decade.

Large increases in the Latino and Asian populations were responsible for the changes in the share white. Over the forty-year period, the Latino population increased from just under ten percent of the population to become a quarter of the population of the large urban areas in 2020. Percentage increases were substantial in each decade but were higher in the middle decades. The share of the population Asian, while much

Table 1. Percent of total population and change in percent for four racial and ethnic groups in 56 large urban areas, 1980-2020

	White	*Black*	*Latino*	*Asian*
Percent of total population				
1980	72.0	15.8	9.5	2.7
1990	66.5	16.0	12.8	4.8
2000	59.9	16.0	17.9	6.3
2010	54.1	15.9	22.1	7.9
2020	49.6	15.7	24.7	9.9
Change in percent of total population				
1980-1990	-5.5	0.2	3.3	2.0
1990-2000	-6.7	0.0	5.1	1.5
2000-2010	-5.7	-0.1	4.2	1.7
2010-2020	-4.6	-0.1	2.7	2.0

smaller, more than tripled from less than three percent to virtually ten percent and increased steadily.

The focus here is on the racial and ethnic diversity of the population at the neighborhood level, in census tracts. As described above, this is measured using the index of diversity. Table 2 has summary statistics for tract diversity and change for all of the tracts in the large urban areas from 1980 to 2020. Mean tract diversity jumps from 32 in 1980 to 59 in 2020. This is obviously a very large increase, but some examples will give an idea of the meaning of this change. The mean index of 32 for 1980 would be the value for a tract with 86 percent of the population in one of the groups, 14 percent in a second group, and none in the other two. The value of 59 in 2020 could represent a tract with 67 percent in one group and 33 percent in the second, or 73 percent in one group and 9 percent in each of the other three groups.

Values for tract diversity encompass the full range from 0, members of only one group present in the tract, to nearly 100, approximately equal proportions for all four groups. The middle half of the tracts in terms of diversity fall in a much narrower range. For 1980, this goes from 10 to about 60. Just as the mean increased dramatically, this range shifts up to 41 to 79 for 2020. The first quartile, the bottom of this range in 2020, is significantly higher than the mean level of diversity of 32 in 1980.

Table 2. Summary statistics for census tract diversity and change tract diversity in 56 large urban areas, 1980-2020

	Mean	*Minimum*	*First quartile*	*Median*	*Third quartile*	*Maximum*
Census tract diversity						
1980	31.6	0.0	10.1	24.1	50.9	99.9
1990	39.3	0.0	15.6	34.9	62.4	99.8
2000	46.7	0.0	23.1	46.3	69.6	99.9
2010	53.7	0.0	32.7	56.6	75.0	99.9
2020	59.3	0.0	41.2	63.6	78.8	99.9
Change in tract dversity						
1980-1990	8.7	-70.5	0.4	7.1	16.5	92.5
1990-2000	8.0	-84.7	0.0	7.1	16.5	89.2
2000-2010	7.2	-92.8	-0.1	7.0	14.7	85.7
2010-2020	5.7	-90.4	0.2	5.6	11.2	84.5

Now for the change in tract diversity from one census to the next, with the statistics in the lower portion of Table 2. Each decade saw substantial mean change in tract diversity. The mean change did decline, from nearly 9 points for the period from 1980 to 1990 to under 6 percent in the final period from 2010 to 2020. Changes in tract diversity across all of the tracts and periods ranged from a drop of over 90 to an increase of over 90, nearly the largest swings possible. But these values reflect very small numbers of extreme outliers. It is more useful to look at the quartiles and the range encompassing the middle half of tract diversity change. The first quartile was very close to zero in each of the four decades. In other words, approximately three-quarters of the tracts saw increases in diversity. The third quartile was over 16 in the first two decades, dropping to around 11 in the final decade. This is consistent with the smaller mean change for the later periods.

Since the focus is on the change in tract diversity, it is useful to take a further look at the distribution of the changes. Figure 1 shows a histogram for the distribution of change in tract diversity for the period from 1980 to 1990. (Histograms for the later periods are similar.) The bars are 5 units wide and the highest bar is for change from 0 to 5, representing 6,036 of 32,092 tracts, 19 percent of all tracts. The heights of the bars decline far more rapidly for smaller changes than for larger, reflecting the fact that about three-quarter of the tracts have positive changes. But numbers drop off quickly in both directions. Ninety percent of the tracts have changes between -13 and 34, and 98 percent

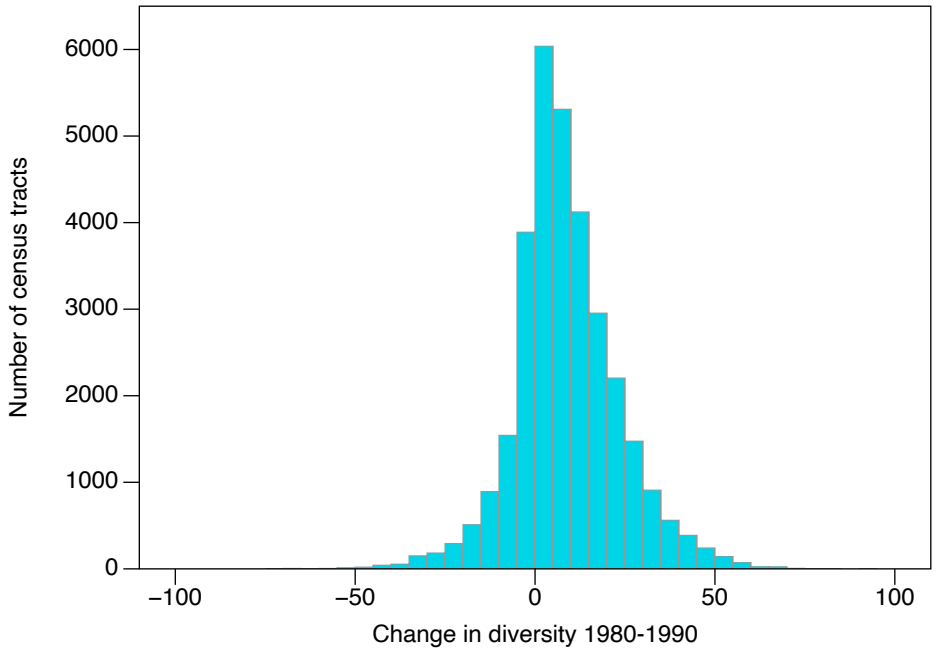

Figure 1. Histogram of change in census tract diversity from 1980-1990

are between -29 and 49. So only the bottom one percent and top one percent of the tracts experienced more extreme changes.

This research examines change in racial and ethnic diversity in the full set of census tracts in the 56 large urban areas (with characteristics of those urban areas included in the analysis). Differences in tract diversity and change exist among the urban areas, and it useful to examine that variation. Table 3 includes basic statistics—mean, minimum, and maximum—for the overall diversity of the urban areas and for the mean diversity and mean change in diversity for the census tracts in each urban area. Because tract populations vary, the latter two measures are the means across the tract values weighted by the tract populations. Values are given only for 1980 and 2020 to streamline the presentation. Intermediate years have intermediate values.

The overall diversity for an urban area is simply the index of diversity calculated using the proportions of the population of the entire urban area in each of the four racial and ethnic groups. Overall diversity increased, which was a prerequisite for the increase in mean tract diversity . The mean of the diversity values for the 56 urban area were 49 in 1980 and 74 in 2020. Because these are values for entire urban areas, variation was not as extreme as for tract diversity, but the range was still substantial, from 14 for the least diverse urban area in 1980 to 77 for the urban area with the most diverse population. The range for 2020 shifted upward to 36 to 93.

Mean tract diversity for each urban area was the weighted mean of the diversity for the tracts in that urban area. The mean of these values increased from 28 to 57, very close to the mean for all tracts reported in Table 2. These values and the minima and maxima are necessarily lower than the corresponding values for urban area overall diversity. This is because the maximum possible mean tract diversity for an urban area

Table 3. Basic statistics for urban area overall diversity and mean census tract diversity and mean change in tract diversity for 56 large urban areas, 1980 and 2020

	Mean	*Minimum*	*Maximum*
Urban area overall diversity			
1980	49.2	14.1	77.0
2020	73.8	36.0	92.9
Urban area mean tract diversity			
1980	28.1	11.6	58.5
2020	57.2	32.8	80.1
Urban area mean change in tract diversity			
1980-2020	28.3	-16.3	46.1

will be its overall diversity, achievable only when every tract has the same distribution across the racial and ethnic groups as the urban area as a whole.

The mean change in tract diversity is calculated in an analogous manner to the weighted means of the change in diversity for the urban area tracts. The mean change from 1980 to 2020 was an increase in mean tract diversity of 28, coincidently the same as mean change in tract diversity across all tracts. This ranged from a decline of 16 points to an increase of 46.

A more detailed view of urban area mean tract diversity and change is provided by looking at the areas with the highest and lowest change. Table 4 gives the 6 urban areas (about 10 percent of the 56) with the greatest change in mean tract diversity and the 6 areas with the smallest change. The areas at the top all experienced mean change in tract diversity greater than 40, large increases given that the diversity index ranges from 0 to 100. This is a varied assortment of urban areas. Minneapolis-St. Paul had the minimum value of mean tract diversity in 1980 and all but one of the others had mean tract diversity below the mean for all urban areas. That the initial diversity values

Table 4. Urban areas with the largest and smallest mean change in census tract diversity from 1980 to 2020 and mean tract diversity in 1980 and 2020

Urban area	*Urban area mean change tract diversity 1980-2020*	*Urban area mean tract diversity*	
		1980	*2020*
Minneapolis-St Paul	46.1	11.6	53.7
Oklahoma City	45.5	21.5	60.5
Seattle-Tacoma	44.3	23.6	67.5
Orlando	44.2	23.3	72.0
Indianapolis	41.5	21.1	54.1
Las Vegas	40.6	36.0	80.1
...	...	...	...
Los Angeles	12.7	50.3	64.7
Memphis	11.0	25.2	50.3
Albuquerque	10.8	53.1	63.8
San Antonio	-0.6	47.6	63.0
Honolulu	-2.8	58.5	56.9
El Paso	-16.3	49.2	32.8

tended to be low makes sense as that provided room for the large increases. By 2020, mean diversity in two areas remained below the mean (but relatively higher) while three of the areas reached levels of mean diversity in the top quarter of all urban areas, with Las Vegas having the highest mean tract diversity of all 56 urban areas.

Of the urban areas with the lowest change in mean tract diversity, three experienced declines from 1980 to 2020, though only El Paso had a substantial decline, -16. The other three areas saw increases greater than 10. The mean tract diversity for each of these areas in 1980 was in the top quarter of all urban areas, with 5 in the top 6. They were more limited in terms of the opportunity for increase than the areas with the highest increases in mean tract diversity. The smaller increases in mean tract diversity dropped the position of these areas relative to the other urban areas in 2020, with 5 of the 6 ranging from somewhat below the mean to Los Angeles, right at the third quartile. El Paso was significantly lower at the first quartile. El Paso frequently emerges as an outlier in these analyses of diversity, as it was 85 percent Latino in 2020. Large populations of groups other than white characterize most of the other areas as well. San Antonio and Albuquerque were the other two urban areas in which the majority of the population was Latino, Honolulu was the one area majority Asian and Pacific Islander, and Memphis the one area majority black.

Factors associated with change in neighborhood diversity

Models are developed predicting the change in tract racial and ethnic diversity from 1980 to 1990 and from 2010 to 2020. This section begins with a discussion of the variables included in the model—both tract and urban area characteristics—thought to be associated with diversity change. The methods used and approaches taken in developing the models are explained. Results are presented and discussed.

First consider tract characteristics potentially associated with the change in tract racial and ethnic diversity. Population growth provides more opportunity for change in the racial and ethnic makeup of a tract and can allow greater increases in diversity. The level of diversity at the start of the period could be related to change in diversity, but with effects in either direction. Tracts with higher levels of diversity could have populations more tolerant and accepting of greater diversity, leading to even higher levels. Alternatively, tracts with low levels of diversity might provide greater opportunity for increases.

Since diversity depends on the percentages of the population in the racial and ethnic groups, change in diversity will necessarily be associated with changes in those percentages. The question is whether a pattern exists between changes in the different groups and changes in diversity. Because whites are the largest group, declines in percent white would allow for and contribute to increased diversity. Likewise, increases

in the percentages black, Latino, and Asian would be expected to be associated with increases in diversity.

The distribution of the population at the start of the period could also affect change in diversity, though in the opposite direction. Greater white population would provide more possibility for decrease in that population and therefore increases in diversity. Conversely, large black, Latino, and Asian populations could mean that diversity is already higher with less scope for increase.

The argument thus far has been that characteristics of the census tracts would be associated with change in diversity. Crowder, Pais, and South (2012) show that metropolitan area characteristics also affect migration of racial and ethnic groups between neighborhoods. This can occur due to variation in housing supply and demand and the opportunities for selecting neighborhoods with different racial compositions. It follows that characteristics of the urban areas may also be related to changes in racial and ethnic diversity at the tract level. Similar arguments to those made for the potential effects of tract characteristics on change can be made for the comparable urban area characteristics.

Shifting now to model development. First, a small point on the population change variable. Two measures of change were tried, percent change over the decade and the change in the log of population. In one year, percent change did slightly better and in the other year, change in the log of population did better. Since the improvement with the latter was somewhat greater, for consistency the change in the log of population was used as the measure of population change in all of the models.

The percentages in the racial and ethnic groups are not independent; they sum to 100. As a result, they are perfectly collinear. When all are included in a model, the model cannot be estimated. The standard response to such a situation is to drop one of the variables. In this case, the best choice would be to exclude whites. Whites are the largest group and if included with the percentages for two of the other groups, problems with multicollinearity can still arise. The problem with excluding whites is that one can only draw inferences on the relationship of the white population to change in diversity from the results for the other three groups, which may vary. But given that whites are the largest group, the relationship to change in diversity is certainly of interest. To address this, two models are estimated for each period, one with the variables for whites only and one with the variables for blacks, Latinos, and Asians.

The units of analysis are the census tracts, which are located within the 56 large urban areas. Tracts within an urban area are subject to similar influences. Urban areas characteristics are repeated for all tracts within each area so degrees of freedom based on the total number of tracts are inappropriate. Error terms are likely to be correlated, violating a basic assumption of ordinary least squares. For these reasons, multilevel model regression is employed to estimate the models, using the *mixed* procedure in Stata. Running the procedure to predict the change in tract diversity in the first and last

decade using an empty model with no predictors produced results that were statistically significantly better than a linear model, with intraclass correlations of 0.082 and 0.074 for the two periods, supporting the use of multilevel estimation.

Random intercept models were estimated predicting the change tract diversity from 1980 to 1990 and from 2010 to 2020 for models including the variables for whites and the variables for blacks, Latinos, and Asian. A step-by-step procedure was used. In the first step, tract population change and diversity were included. The change in the tract percent in the population groups were then added, whites in one case and blacks, Latinos, and Asians in the other. This was followed by addition of the percentage(s) in the same group(s). Three more steps added the corresponding urban area characteristics to the model. For all but two additions, the likelihood ratio test indicated that the latter model was significantly better ($p < 0.05$) and and the two instances of nonsignificance occurred in different years and at different stages in the procedure. In most cases, addition of new variables resulted in minimal to modest changes in the regression coefficients previously added to the model. In only a few instances were levels of significance affected.

I had no special reason to expect coefficients to vary across urban areas. But on an exploratory basis, random slope models were estimated using the change in log of tract population and tract diversity, the two tract measures common to the various models. The effects compared to the random slope models varied and were inconsistent across the models for whites versus blacks, Latinos, and Asians and across the models for 1980-1990 versus 2010-2020. Some of the changed results were illogical compared to the random slope results. The conclusion was that the random slope models were not useful.

A final issue involves the measurement of goodness-of-fit. Multilevel models do not have the traditional R^2 measure familiar from ordinary least squares. Efforts have been made to devise pseudo-R^2 measures of goodness of fit at the different model levels, in the present case census tracts and urban areas. Stata was used to produce two sets of measures proposed by different pairs of authors: Snijders/Bosker R^2 and Bryk/Raudenbush R^2. An alternative, simple approach was taken to measure the fit of the models. The values of change in tract diversity predicted by each model were regressed on actual change in tract diversity, producing a value of R^2 that is a measure of the proportion of the variance accounted for by that prediction. These predicted R^2 values differed little from the Snijder/Bosker level 1 R^2 values, varying only either in the third significant digit or by one in the second significant digit. The values produced using the predicted values and regression, called here the *Prediction* R^2, are reported. An entry in Stata Frequently Asked Questions by Cox (2024) suggests the use of the equivalent procedure of squaring the correlation of the predicted and actual values.

Table 5 presents the basic multilevel model results including the unstandardized regression coefficients and some summary measures. The first two columns of results

are for models predicting change in tract diversity from 1980 to 1990. The final columns are for change from 2010 to 2020. Results for two models are reported for each period. The first includes the variables for whites, and the second the variables for blacks, Latinos, and Hispanics, as discussed above. (Models were also estimated for the intermediate periods, but since the results fall between those presented here for the first and last period, they add little and are not reported.)

The estimated regression coefficients for the fixed effects are presented in the top two sections. They are divided to make clear the distinction between the census tract characteristics and the urban areas characteristics. Next are the variances for the random effects at the tract and urban area levels. The last section includes summary measures for the whole model.

The intraclass correlations coefficients are very small, ranging from 0.026 to 0.044. This means that only a very small proportion of the variation is accounted for by the level 2 urban area characteristics. These values are much smaller than the intraclass correlation coefficients for the null model with no fixed effects which were 0.082 and 0.074. A significant portion of the variance associated with the urban areas in the null models is due to variation in the tract characteristic across the urban areas and is now accounted for in these models by the tract characteristics.

The prediction R^2 value for the models range from 0.25 to 0.45, so the predicted values are accounting for moderate proportions of the variation in the change in tract diversity. In each period, R^2 for the second model with the variables for blacks, Latinos, and Asians is higher than the model with the variables for whites. This makes sense, as having the variables for the three groups provides more information and the opportunity for a better fit than only having the information on the white (or nonwhite) population. The R^2 values for the models predicting change from 1980 to 1990 are noticeably larger than those for 2010 to 2020. This may be associated with the lower average increases in diversity in the latter period.

The following procedure will be followed in discussing the large numbers of regression coefficients. The initial focus is on the coefficients for the tract characteristics in the upper half of the table for the earlier period from 1980 to 1990. These will then be compared with the tract characteristics for the later decade. Consideration of the coefficients for the urban area characteristics follows.

All of the tract coefficients for the two 1980-1990 models are statistically significant with *p*-values less than 0.001. Higher tract population growth as measured by the change in the logarithm of tract population is associated with increased levels of diversity, as expected. The coefficient for tract diversity at the beginning of the decade is negative. The earlier discussion of the possible effect of diversity suggested possible effects in both directions. At the tract level, it appears that lower levels of diversity may be providing greater opportunities for increases in diversity.

Table 5. Models predicting change in tract diversity estimated using multilevel regression models with random intercepts

Independent variables	*Change diversity 1980-1990*		*Change diversity 2010-2020*	
Change log tract population	1.571 ***	1.403 ***	3.041 ***	2.471 ***
Census tract diversity	-0.189 ***	-0.168 ***	-0.130 ***	-0.111 ***
Change tract percent white	-0.723 ***	—	-0.297 ***	—
Change tract percent black	—	0.555 ***	—	0.221 ***
Change tract percent Latino	—	0.767 ***	—	0.249 ***
Change tract percent Asian	—	1.400 ***	—	0.684 ***
Tract percent white	0.112 ***	—	0.049 ***	—
Tract percent black	—	-0.084 ***	—	-0.003
Tract percent Latino	—	-0.213 ***	—	-0.126 ***
Tract percent Asian	—	-0.151 ***	—	-0.144 ***
Change log urban area pop	11.577 ***	13.042 ***	2.458	3.447
Urban area overall diversity	0.129 **	0.220 ***	0.086 ***	0.046 *
Change urban area pct white	-0.350 **	—	-0.274 *	—
Change urban area pct black	—	0.287	—	0.287
Change urban area pct Latino	—	-0.183	—	0.325 *
Change urban area pct Asian	—	0.135	—	0.210
Urban area percent white	0.072	—	0.081 ***	—
Urban area percent black	—	-0.245 **	—	-0.067 *
Urban area percent Latino	—	-0.100	—	0.006
Urban area percent Asian	—	-0.051	—	0.045
Constant	-15.295 *	1.852	-3.914	8.884 ***
Level 1 (tract) variance	119.510	109.601	70.219	63.794
Level 2 (urban area) variance	5.533	3.135	2.025	1.705
Log likelihood	-122,185	-120,786	-144,107	-142,155
Intraclass correlation coefficient	0.044	0.028	0.028	0.026
Prediction R^2	0.396	0.454	0.245	0.315
N	32,043	32,043	40,632	40,632

** $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$*

Change in the percent of the tract population white is negatively related to the change in diversity. The results for the second model show the corresponding effects of changes in the percent black, Latino, and Asian, with increases in the numbers in those groups related to growth in diversity. The coefficients for the percentages in the groups show the opposite side of this. A larger share of the population white provides greater possibilities for decline and is thus positively related to change in diversity. This is also necessarily related to lower proportions black, Latino, and Asian, hence the negative coefficients.

Shifting attention to the tract variable coefficients for 2010 to 2020, all but one have high statistical significance and all have the same signs as the corresponding variables in the earlier period. Except for the change in the log of tract population, the 2010-2020 regression coefficients are nearly all smaller than their earlier counterparts, some substantially so. The mean change in tract diversity in the final decade is 5.7, significantly smaller than the change of 8.7 in the first period. This may be associated with the declines in the regression coefficients. The exceptions are the coefficients for the population change variable which increases from 1.6 and 1.4 in the earlier models to 3.0 and 2.5 for 2010-2020. I have no idea why this increase might have occurred.

Shifting to the coefficients for the urban area characteristics, the initial observation is that only half of the coefficients for the four models are statistically significant, even at the 0.05 level, and many of those have lower levels of significance than the uniformly high levels for the tract variables. This is reasonable given less variation across these variables at the tract level (the same for all tracts in an urban area) and the low proportion of the variance accounted for by the urban area variables compared to the tract variables.

The coefficients for urban area population change from 1980 to 1990 are 11.6 and 13.0, very much larger than the tract coefficients and the urban area coefficients for final period, 2.5 and 3.4. This variable continues to confound, and I have no ideas about why this might have occurred either. The coefficients for urban area diversity are interesting: All are positive and statistically significant, the opposite from tract diversity. The alternative idea proposed for this variables is that greater diversity would be associated with a more tolerant population more likely to lead to further increases in diversity. It seems that this could be the case, with diversity at the urban area level having this effect opposite the tract level effect.

The change in urban area percent white is negatively related to change in diversity, the same as at the tract level, though with weaker levels of significance. Five of the six among percent black, Latino, and, and Asian in the two decades are positive but only one is statistically significant, so relationships to tract change are greatly diminished compared to the tract level. That one has a negative sign should not be considered meaningful as the coefficient was not statistically significant. Urban area percent white is positive in both periods but statistically significant only in the latter.

Urban area percent black in negative and significant in the two decades as well. Coefficients for percent Latino and Asian are nonsignificant and very small, so the presence of two with positive signs is not surprising.

The unstandardized coefficients for the tract variables are have signs consistent with expectations and high levels of statistical significance. With a few exceptions, the regression coefficients for the urban area characteristics are smaller than the tract-level counterparts and less likely to be statistically significant. The more interesting result for these is the difference in the directions of the relationships of tract diversity and urban area overall diversity to change in tract diversity.

The unstandardized regression coefficients allow comparisons between the results for the two years for each variable as well as comparison of tract and urban area coefficients for the corresponding variables. These coefficients are estimates of the relationship of a one-unit change in each variable to change in tract diversity. Because variables are measured in different units, the unstandardized coefficients cannot be compared across variables to assess the relative effects of the variables on change in diversity. Lorah (2018) discusses the assessment of effect sizes in multilevel models and suggests the use of standardized regression coefficients for the fixed effects. Standardized regression coefficients represent the change in the number of standard deviations in the dependent variable, change in tract diversity, associated with a one standard deviation change in an independent variable. The magnitudes of the standardized regression coefficients can be interpreted as measures of the relative effects of the independent variables on the dependent variable. Standardized regression coefficients have been calculated by multiplying the unstandardized coefficients by the ratio of the standard deviation in tract diversity to the standard deviation of each of the variables. The results are presented in Table 6.

The magnitudes of the standardized regression coefficients are the measures of relative importance. The signs indicate the direction of the relationship and are the same as the signs of the unstandardized coefficients Considering the coefficients in Table 6, to focus the discussion, coefficients with magnitudes greater than 0.3 are considered to represent a large effect, coefficients between 0.1 and 0.3 a moderate effect, and smaller coefficients suggest small effects diminishing to minimal. Beginning with the two models predicting change in tract diversity from 1980 to 1990, the large effects are associated with tract characteristics. Change in tract percent white, black, Latino, and Asian have large effects as does tract diversity in the first model and nearly so in the second, 0.29. Tract percent white, black, and Latino have moderate effects between 0.15 and 0.24. The remaining tract characteristics, change in log of population and percent Asian have small effects with the former perhaps considered minimal.

No urban area variables have large effects. Change in the log of urban area population and urban area overall diversity have moderate effects. Urban area percent

Table 6. Standardized regression coefficients for models predicting change in tract diversity

Independent variables	*Change diversity 1980-1990*		*Change diversity 2010-2020*	
Change log tract population	0.036	0.032	0.114	0.093
Census tract diversity	-0.331	-0.294	-0.336	-0.287
Change tract percent white	-0.511	—	-0.226	—
Change tract percent black	—	0.307	—	0.132
Change tract percent Latino	—	0.365	—	0.147
Change tract percent Asian	—	0.391	—	0.288
Tract percent white	0.237	—	0.155	—
Tract percent black	—	-0.153	—	-0.006
Tract percent Latino	—	-0.243	—	-0.301
Tract percent Asian	—	-0.074	—	-0.174
Change log urban area pop	0.115	0.130	0.023	0.033
Urban area overall diversity	0.147	0.252	0.124	0.066
Change urban area pct white	-0.072	—	-0.051	—
Change urban area pct black	—	0.023	—	0.030
Change urban area pct	—	-0.027	—	0.044
Change urban area pct Asian	—	0.011	—	0.023
Urban area percent white	0.068	—	0.128	—
Urban area percent black	—	-0.177	—	-0.077
Urban area percent Latino	—	-0.083	—	0.009
Urban area percent Asian	—	-0.029	—	0.043

black also has a moderate effect in its model. Effects of other variables are relatively small with some of those minimal.

Shifting to the models predicting change from 2010 to 2010, tract diversity in the first model has a large effect and the coefficient in the second model again comes close. The tract percent Latino also has a large effect, somewhat greater than the substantial moderate effect for 1980 to 1990. All of the variables for change in tract percent in one of the four groups shift from large effects in the earlier year to moderate effects in the latter period. In addition to the large effect for percent Latino, tract percent white and percent Asian also have moderate effects. Overall urban area diversity in the first model and urban area percent white are the only urban area variable with moderate effects, just

above the threshold for that category. All of the other urban areas characteristics have small to very small effects on change in tract diversity.

Comparing 2010-2020 standardized regression coefficients with those for the earlier year, change in the log of tract population has relative greater influence on change in diversity, and tract diversity has a similar effect. The variables for the change in percent and percent in the four groups show a generally similar pattern to those for the earlier period though values tend to be smaller. The effects for the urban area characteristics are either smaller for the latter period or the values are so small that differences are irrelevant.

Conclusions

Three of the findings of this research are highlighted in the conclusion. Tract racial and ethnic diversity increased greatly. The differing relationships of tract and urban area diversity to the change in diversity proved interesting. And the relationship of change in diversity to the changes in the shares of the population in the racial and ethnic groups raises issues for further consideration.

The increase in the racial and ethnic diversity of the census tracts in the 56 large urban areas from 1980 to 2020 was very large. Not only did the mean jump from 32 to 59, the entire distribution moved up, with the first quartile diversity in 2020 far higher than the 1980 median. About three quarters of all tracts saw increases in diversity in each decade.

A major factor contributing to this increase in tract diversity was the increase in the overall diversity of the urban areas over this period. Mean tract diversity in an urban area can be no higher than the urban area overall diversity (Ottensmann 2023). In 40 of the 56 urban areas, mean tract diversity in 2020 was higher than urban area diversity in 1980. Thus increases in urban area diversity were necessarily required to allow for the observed increases in tract diversity.

Increases in overall diversity do not require than tract diversity increase. Highly segregated residential patterns could be established and tract diversity could even decline. While we know that the increase in urban area overall diversity was required to allow for increased diversity at the tract level, we cannot know from these results the extent to which increased tolerance for living in diverse neighborhoods may have also contributed.

The discussion of possible relationships of various tract and urban area characteristics to change in tract diversity speculated that diversity at the start of the period could have effects in either direction. Greater diversity might reflect higher levels of tolerance, allowing greater increases in tract diversity. Alternative, low levels of diversity might provided greater opportunity for increases in tract diversity. Results for the models predicting change in tract diversity showed opposite relationships for tract

diversity and urban area diversity. The diversity of the tract at the start of the period was negatively related to the change in tract diversity. This is consistent with the hypothesis that lower levels of diversity make larger increases possible. On the other hand, urban area overall diversity at the start was positively related to the change in tract diversity, potentially reflecting the effect of higher levels of tolerance at the urban area level. Some of the surveys of individual preferences for and willingness to live in racially mixed neighborhoods have suggested increases in tolerance over time. The positive relationship of urban area diversity to change in diversity is consistent with this, though certainly cannot be taken as clear evidence in support of such increase.

The increase in tract diversity associated with the decline in the share of whites in a tract is at the same time obvious but interesting. Obvious because more neighborhoods have relatively high shares of whites so that a decrease along with increased numbers from other racial and ethnic groups would result in more diversity. However, much of the research on racial change has focused on blacks and whites and on racial succession, the tipping point, and then later the possibility of stable integrated neighborhoods. The dominant idea was that in racially mixed neighborhoods, the share of whites would decline either as a result of white flight or by whites refusing to move into such neighborhoods, resulting in transition to mainly black neighborhoods. But the finding here is that the decline in percent white is associated with increase in diversity, just the opposite. So is the increase in diversity the decline in the share white or the increase in the shares of other groups? If tract population remains constant (or relatively so), there is no way to distinguish the effects.

The increase in urban area overall diversity is the result of increases in the numbers and shares of the population Latino and Asian, with little change in the share black. This raises the question of exactly how this is affecting the change in tract diversity. The results have shown that not only are increases in the shares of tract populations Latino and Asians associated with increases in diversity, increase in the share black are as well. A lot more remains to be untangled in the understanding of change in the racial and ethnic composition of neighborhoods.

References

Alba, Richard D., Nancy A. Denton, Shu-yin J. Leung, and John R. Logan. 1995. Neighborhood change under conditions of mass immigration: the New York City region, 1970–1990. *International Migration Review* 29: 625–656.

Aldrich, Howard. 1975. Ecological succession in racially changing neighborhoods. *Urban Affairs Review* 10, 3: 327-348.

Bobo, Lawrence, and Camille L. Zubrinsky. 1996. Attitudes on residential integration: perceived status differences, mere in-group preferences, or racial prejudice. *Social Forces* 74: 883-909.

Brown, Lawrence A., and Madhuri Sharma. 2010. Metropolitan context and racial/ethnic intermixing in residential space: U.S. Metropolitan Statistical Areas 1990-2000. *Urban Geography* 31: 1-28.

Budescu, David V., and Mia Budescu. 2012. How to measure diversity when you must. *Psychological Methods* 17: 215-227.

Burgess, Ernest W. 1928. Residential segregation in American cities. *Annals of the American Academy of Political and Social Science* 140: 105-115.

Card, David , Alexandre Mas, and Jesse Rothstein. 2008. Tipping and the dynamics of segregation. *Quarterly Journal of Economics* 123: 177-218.

Charles, Camille Zubrinsky. 2000. Neighborhood racial-composition preferences: evidence from a multiethnic metropolis. *Social Problems* 47: 379-407.

Clark, William A.V. 1992. Residential preferences and residential choices in a multiethnic context. *Demography* 29: 451-466.

Cohn, D'Vera. 2010. Census history: counting Hispanics. Pew Research Center. Downloaded from *http://www.pewsocialtrends.org/2010/03/03/census-history-counting-hispanics-2/* on March 14, 2019.

Cox, Nicholas J. 2024. How can I get an R-squared value when a Stata command does not supply one? Downloaded from *https://www.stata.com/support/faqs/statistics/r-squared/* on February 28, 2024.

Crowder, Kyle, Jeremy Pais, and Scott J. South. 2012. Neighborhood diversity, metropolitan constraints, and household migration. *American Sociological Review* 77: 325-353.

Denton, Nancy A., and Douglas S. Massey. 1991. Patterns of neighborhood transition in a multiethnic world: U.S. metropolitan areas, 1970-1980. *Demography* 28: 41-63.

Duncan, Otis Dudley, and Beverly Duncan. 1957. *The Negro Population of Chicago: A Study of Residential Succession*. University of Chicago Press.

Easterly, William. 2009. Empirics of strategic interdependence: the case of the racial tipping point. *B.E. Journal of Macroeconomics* 9, 1 (Contributions), Article 25.

Ellen, Ingrid Gould. 1988. Stable racial integration in the contemporary United States: an empirical overview. *Journal of Urban Affairs* 20: 27-42.

Ellen, Ingrid Gould. 2000. *Sharing America's Neighborhoods: The Prospects for Stable Racial Integration*. Harvard University Press.

Farley, Reynolds, Howard Schuman, Suzanne Bianchi, Diane Colasanto, and Shirley Hatchett. 1978. Chocolate city, vanilla suburbs: will the trend toward racially separate communities continue? *Social Science Research* 7: 319-344.

Farley, Reynolds, Elaine L. Fielding, and Maria Krysan. 1997. The residential preferences of whites and blacks: a four-metropolis analysis. *Housing Policy Debate* 8: 763-800.

Farrell, Chad R. 2005. *Urban Mosaics: Multiracial Diversity and Segregation in the American Metropolis*. Ph.D. thesis, Pennsylvania State University.

Farrell, Chad R. 2008. Bifurcation, fragmentation, or integration? The racial and geographic structure of metropolitan segregation, 1990-2000. *Urban Studies* 45: 467–499.

Farrell, Chad R., and Barrett A. Lee. 2011. Racial diversity and change in metropolitan neighborhoods. *Social Science Research* 40: 1108-1123.

Friedman, Samantha. 2008. Do declines in residential segregation mean stable neighborhood racial integration in metropolitan America? A research note. *Social Science Research* 37: 920-933.

Galster, George. 1998. A stock/flow model of defining racially integrated neighborhoods. *Journal of Urban Affairs* 20: 43-51.
Goering, John M. 1978. Neighborhood tipping and racial transition: a review of social science evidence. *Journal of the American Institute of Planners* 44: 68-78.
Grodzins, Morton. 1957. Metropolitan segregation. *Scientific American* 197, 4 (October): 33-41.
Grodzins, Morton. 1958. *The Metropolitan Area as a Racial Problem*. University of Pittsburgh Press.
Herfindahl, Orris C. 1950. *Concentration in the U.S. Steel Industry*. Ph.D. dissertation, Columbia University.
Hipp, John R., and Jae Hong Kim. 2023. Persistent racial diversity in neighborhoods: what explains it and what are the long-term consequences. *Urban Geography* 44: 640-667.
Hirschman, Albert O. 1964. The paternity of an index. *American Economic Review* 54: 761-762.
Iceland, John. 2004. Beyond black and white metropolitan residential segregation in multi-ethnic America. *Social Science Research* 33: 248-271.
Lee, Barrett A., and Peter B. Wood. 1991. Is neighborhood racial succession place-specific? *Demography* 28: 21–40.
Logan, John R., and Charles Zhang. 2010. Global neighborhoods: new pathways to diversity and separation. *American Journal of Sociology* 115: 1069–1109.
Lorah, Julie. 2018. Effect size measures for multilevel models: definition, interpretation, and TIMSS example. *Large-scale Assessments in Education* 6, 8.
Maly, Michael T. 2000. The neighborhood diversity index: a complementary measure of racial residential settlement. *Journal of Urban Affairs* 22: 37-47.
Maly, Michael T. 2005. *Beyond Segregation: Multiracial and Multiethnic Neighborhoods in the United States*. Temple University Press.
Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. 2022. *IPUMS National Historical Geographic Information System: Version 17.0* [dataset]. Minneapolis MN: IPUMS. *http://doi.org/10.18128/D050.V17.0*. Data downloaded from *https://www.nhgis.org* on various dates, August to October 2022.
Massey, Douglas S. 1983. A research note on residential succession: the Hispanic case. *Social Forces* 61: 825-833.
Massey, Douglas S., and Nancy A. Denton. 1987. Trends in the residential segregation of Blacks, Hispanics, and Asians. *American Sociological Review* 52: 802-825.
Ottensmann, John R. 2023a. Urban patterns of the largest urban areas in the United States, 1950-2020: a new dataset. Available at *https://urbanpatternsblog.files.wordpress.com/2023/02/new-urban-patterns-dataset.pdf*.
Ottensmann, John R, 2023b. Racial and ethnic diversity in neighborhoods in large urban areas in the U.S., 1980-2020. Available at *https://urbanpatternsblog.files.wordpress.com/2023/12/neighborhood-racial-ethnic-diversity.pdf*.
Ottensmann, John R., and Michael E. Gleason. 1992. The movement of whites and blacks into racially mixed neighborhoods: Chicago, 1960-1980. *Social Science Quarterly* 73: 645-662.
Park, Robert E. 1936. Succession, an ecological concept. *American Sociological Review* 1: 171–179.
Pew Research Center. 2013. Preference for the terms "Hispanic" and "Latino." Downloaded from *https://www.pewhispanic.org/2013/10/22/three-fourths-of-hispanics-say-their-community-needs-a-leader/ph-hispanic-leader-10-2013-03-04/* on June 20, 2019.
Rapkin, Chester, and William G. Grigsby. 1960. *The Demand for Housing in Racially Mixed Areas*. University of California Press.

Rawlings, Lynette, Laura Harris, Margery Austin Turner, with Sandra Padilla. 2004. Race and residence: prospects for stable neighborhood integration. *Neighborhood Change in Urban America*. Urban Institute, no. 3 (March). Downloaded from *https://www.urban.org/research/publication/race-and-residence* on January 19, 2024.

Reardon, Sean F., and Glenn Firebaugh. 2002. Measures of multigroup segregation. *Sociological Methodology* 32: 33-67.

Schelling, Thomas C. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1: 143-186.

Shannon, Claude E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27: 379-423.

Sigelman, Lee, Steven A. Tuch, and Jack K. Martin. 2005. What's in a name? Preferences for "black" versus "African-American" among Americans of African descent. *Public Opinion Quarterly* 69: 429-438.

Simpson, E. H. 1949. Measurement of diversity. *Nature* 163 (April 30): 688.

Smith, Richard A. 1998. Discovering stable racial integration. *Journal of Urban Affairs* 20: 1-25.

Taeuber, Karl E., and Alma F. Taeuber. 1965. *Negroes in Cities: Residential Segregation and Neighborhood Change*. Aldine.

Taub, Richard P., D. Garth Taylor, and Jan D. Dunham. 1984. *Paths of Neighborhood Change: Race and Crime in Urban America*. University of Chicago Press

Theil, Henri. 1972. *Statistical Decomposition Analysis*. North Holland.

U.S. Census Bureau. 2022. Urban Area criteria for the 2020 Census—final criteria. *Federal Register* 97, 57 (March 24): 16706-16715.

Urban Institute and Geolytics. 2003. *Census CD Neighborhood Change Database (NCDB): 1970-2000 Tract Data*. Geolytics, Somerville, NJ.

White, Michael J. 1986. Segregation and diversity: measures in population distribution. *Population Index* 52: 198-221.

Wolfe, Eleanor P. 1963. The tipping-point in racially changing neighborhoods. *Journal of the American Institute of Planners* 29: 217-222.

Wood, Peter B., and Barrett A. Lee. 1991. Is neighborhood racial succession inevitable? Forty years of evidence. *Urban Affairs Quarterly* 26: 610-620.