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
December 2023

Abstract

Racial and ethnic residential diversity in urban areas is measured using an index of exposure diversity, which is the mean of diversity within census tracts, and an index of evenness diversity relative to the diversity of the urban area. These indexes are used to examine neighborhood diversity among whites, blacks, Latinos, and Asians in 56 large urban areas in the United States from 1980 to 2020. Mean diversity across the urban areas increased dramatically. But substantial differences exist between exposure diversity and evenness diversity. An area with very low exposure diversity can have very high evenness diversity when the overall diversity of the urban area is also low. Diversity tends to be lower in urban areas in the Northeast and Midwest, highest in the West. Lower percent population white and decreases in the white population in the prior decade are associated with higher levels of exposure diversity with just the opposite for blacks, Latinos, and Asians. Average levels of the two measures of diversity increased in both the urban core and suburban periphery, but exposure diversity was higher in the core while evenness diversity was higher in the suburbs.

Introduction

The level of diversity of an urban area describes the mix of racial and ethnic groups in the population—whites, blacks, Latinos, Asians—ranging from low diversity when the population is concentrated in a single group to higher diversity with larger numbers in all groups. Dramatic increases in diversity in large urban areas in recent decades is demonstrated in an earlier paper (Ottensmann 2023c). But this provides little information about the diversity of the populations in individual neighborhoods. If the urban area diversity is low, diversity within neighborhoods is limited by the low numbers of members of the groups that are not dominant. But high levels of diversity in an urban area do not guarantee high neighborhood diversity. Many neighborhoods could still be occupied mainly be members of a single group.

The question of the mix of racial and ethnic groups within neighborhoods has been posed and addressed in the opposite direction, looking at the degree of residential segregation based on the distributions of racial and ethnic populations within and across neighborhoods. This is essentially the opposite of neighborhood level diversity. A neighborhood occupied only by members of a single group is considered to have the highest possible level of segregation and the lowest possible level of diversity. This paper addresses neighborhood diversity in comparison to the indexes used for the measurement of segregation.

The middle of the last century saw the explosion of a huge literature on residential segregation. For a long time, the focus was primarily on segregation between blacks and whites, then the major racial groups in most of the nation's cities. No attempt will be made to review this literature. Two landmarks are the books by Taueber and Taeuber, *Negroes in Cities* (1965), and Massey and Denton, *American Apartheid* (1993). Another study shows the difference in thinking about race and ethnicity from the present. Lieberson (1969) examines residential segregation for ethnic groups, presenting measures of segregation for European foreign-born populations from different countries and their children, illustrating how ideas about ethnic diversity differed from the present.

The transition from the focus on black-white residential segregation to the inclusion of other racial and ethnic groups is nicely illustrated by three articles in a journal issue dealing with the existence of stable racial integration. 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.

As the initial studies of residential segregation focused on two groups, blacks and whites, the indexes used were measures of segregation between two groups. With the growth of other racial and ethnic groups in the population, the first efforts examining their residential segregation used these same measures. Massey and Denton (1987) examine the segregation of whites, blacks, Latinos, and Asians, calculating indexes of segregation among all possible pairs of the groups. Frey and Farley (1996) compute the segregation between each of the groups other than whites and all others. Massey and Tannen (2018) attempt to deal with the large numbers of results by taking weighted averages of the indexes.

In place of the cumbersome process of using multiple indexes of segregation between two groups, measures of multi-group diversity and segregation have been employed in studies of the distribution of racial and ethnic groups within urban neighborhoods. Iceland (2004), Farrell (2008) and Brown and Sharma (2010) examine residential segregation among four or five of the major racial and ethnic groups identified by the Census. Farrell and Lee (2011) take the opposite approach, examining levels of diversity using a related but somewhat different measure.

Two indexes reflecting differing conceptions of diversity (and segregation) are used examine racial and ethnic patterns in large urban areas in the United States from 1980 to 2020. A major focus will be the contrast between these measures. The next two sections describe the data—the urban patterns data providing the context and the data for the racial and ethnic populations in census tracts. This is followed by a discussion of the measurement of residential diversity and segregation and the presentation of the two indexes used. The first set of results address residential diversity across the entire urban areas, how this has changed, how this varies across urban areas, and factors associated with levels of diversity. Diversity in the urban core and the suburban periphery are then compared.

The *Urban Patterns 2* data

The *Urban patterns 2* dataset includes housing unit counts for census tracts from 1950 to 2020 that have been 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 included 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.

Studies of urban diversity have frequently considered the distribution of members of racial and ethnic groups in the central city and the remainder of the urban area, however defined. The latter would be referred to as the "ring" or more often the "suburbs." The basic notion is that the central city encompasses the older portion of the urban area while the suburban ring is the newer, more recently developed area. The problem is that cities are political entities and their boundaries encompass widely varying portions of their urban areas. In some urban areas, territory that had been developed and became a part of the urban area in the nineteenth and early twentieth century are outside the central city, in the area considered to be the suburbs. Conversely, central cities in other urban areas have been able to continue to expand their boundaries such that they include a large share of much more recent development.

The dataset used here allows for the designation of the older and the more recently developed portions of the urban area in a consistent manner. The extent of the urban area in 1940 is delineated using estimates of housing units by census tracts. This is referred to as the urban core. Starting with 1950, that portion of the urban area outside the urban core is considered to be the suburban periphery. Thus the suburban periphery is approximately that part of the urban area developed after World War II, the area often considered to be the suburbs. This urban core-suburban periphery distinction has been used to examine change in the large urban areas (Ottensmann 2023b).

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 cannot 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 group 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 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 discussions 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 white 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 conditions of constituting more than half or less than half of the population.

The measurement of neighborhood diversity and segregation

This section begins with a general discussion of measures of diversity and the presentation of the index to be used in a way that makes clear how it serves as a measure of diversity. This provides the foundation for considering residential diversity and segregation. Exposure and evenness are two dimensions of segregation that have been identified and form the basis for the indexes used in this paper, which are then discussed.

Measures of diversity

Three of the more useful reviews of measures of diversity are White (1986), Reardon and Firebaugh (2002) and Budescue and Budescue (2012). The first two extend their discussions to the measurement of segregation as well. 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

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

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) proposed the index as a measure of species diversity in an ecosystem. Herfindahl (1950) and Hirschman (1964) used the index as a measure of industrial concentration. The formula for the interaction index I is

$$I = 1 - \sum_{i=1}^{n} p_i{}^2 \qquad (2)$$

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

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,

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

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

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

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 adjusted to vary between zero and 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.

Measures of residential diversity and segregation

Measures of residential diversity and segregation within neighborhoods can be seen as functions of the levels of diversity within the census tracts in a larger area, such as the entire urban area. An area is considered to have zero diversity and maximum segregation when every tract is occupied by the members of a single racial and ethnic group. If a measure of residential diversity ranges from zero to one, then the level of segregation is one minus the level of diversity.

Since the middle of the last century a great deal of attention has been devoted to the measurement of residential segregation, especially segregation between two population groups, given the focus on the segregation between blacks and whites. Many indexes have been proposed. This provides a context for considering racial and ethnic diversity and segregation within neighborhoods. Papers by Massey and Denton (1988), White (1986), and Reardon and Firebaugh (2002) provide the foundation for the discussion that follows.

A useful starting point is Massey and Denton, positing five dimensions of residential segregation, describing 20 indexes related to these, and performing an empirical analysis of the interrelationships. Two of the dimensions—exposure and evenness—account for the largest share of the variation among the dimensions and include the indexes most often employed in studies of residential segregation, including by Massey and Denton (1987). These dimension serve as the basis for the measures of residential diversity used in this research.

Exposure measures

Exposure indexes go back to the earliest days of measurement of black-white segregation. The isolation index is the probability of encountering another member of the same group within one's neighborhood and the interaction index is the probability of encountering a member of the other group. The interaction index of diversity, which is the probability that two persons chosen at random from a population will be members of different groups, is the clear extension of this idea to more than two groups. This forms the basis for the first index of residential diversity. Let D_j be the racial and ethnic diversity of census tract j, calculated for each tract using the formula for diversity D in equation (4) above. Then the exposure index of diversity is the mean of the tract diversity weighted by the populations of each census tract:

$$D_{exp} = \sum_{j=1}^{m} \frac{T_j}{T} D_j \qquad (5)$$

where D_{exp} is the exposure diversity for the larger area, T_j is the population of tract j and T is the total population of the larger area. The level of exposure diversity is limited by the level of diversity in the urban area. The maximum possible value of exposure diversity D_{exp} is the overall level of diversity in the urban area and occurs when each tract has the same proportion in each group as the group's proportion in the total urban area population.[1]

White (1986) presents a measure of segregation that is essentially the exposure diversity index presented above subtracted from one to give a measure of segregation. Farrell and Lee (2011) use a measure for exposure diversity that is based on the use of the entropy index as the measure of tract diversity.

Evenness measures

Evenness refers to the distribution of the groups across the census tracts or other subareas. Complete evenness and thus minimum segregation occurs when, for each tract, group shares of their total populations in the urban area are equal. This implies that the percentage distributions across the groups within each tract are equal to each other and to the distribution in the larger area as a whole. The most important of the two-group segregation indexes is the index of dissimilarity, used in far more studies of residential segregation than any other measure. Duncan and Duncan (1955) compare multiple segregation indexes. They suggest that dissimilarity and some of the other indexes are preferable given that their values did not depend on the shares of the population white and black in the urban areas. Whether making such an *a priori* assumption of independence is appropriate is debatable, however. At issue is whether an urban area that is five percent black and 95 percent white with the same distribution in each tract can be considered to have as low a level of segregation as an area 40 percent black and 60 percent white, also with that same distribution in each tract.

Evenness implies the percentage distributions of the racial ethnic groups in all tracts is the same as in the urban area as a whole, meaning that exposure diversity is at the maximum value. An evenness measure of diversity is thus the level of exposure diversity relative to its maximum value, the diversity of the urban area. Evenness diversity D_{even} is thus defined as exposure diversity divided by urban area diversity,

$$D_{even} = \frac{D_{exp}}{D_A} \qquad (6)$$

[1]A simplified demonstration: Starting with proportions in each tract equal to their proportions in the urban area, increasing the proportion in one group in a tract requires an offsetting decrease for another group and changes in the opposite directions in another tract. These higher and lower values are squared in calculating the tract diversities. The increases in the squared values for the larger proportions are greater than the decreases in the squared values for the smaller proportions. The sums of the squares for the changed tracts will be thus be greater, reducing the diversity for these tracts, making the neighborhood diversity less than the urban area diversity.

where D_A is the level of diversity in the urban area or other larger area for which neighborhood-level diversity is being determined. Evenness diversity has the maximum value of one when exposure diversity is has the maximum value equal to urban area diversity. It is zero when exposure diversity is also zero, no members of two or more groups occupying any census tract.

Reardon and Firebaugh (2002) present the relative diversity index of segregation calculated using interaction diversity in a comparable manner. Theil's (1972) information theory index is an analogous measure of segregation using entropy as the measure of tract diversity. It is the weighted index of tract entropy diversity divided by the entropy diversity of the urban area, subtracted from one to produce a measure of segregation. Iceland (2004), Farrell (2008) and Brown and Sharma (2010) use this in studies of metropolitan area racial and ethnic segregation.

In this research the values for the diversity indexes are multiplied by one hundred for presentation. Minimum diversity, the population concentrated in a single group, still has the value of zero and the maximum values for urban area diversity and evenness diversity are one hundred.

Exposure and evenness diversity in large urban areas

This section uses the exposure and evenness diversity indexes to examine levels of residential racial and ethnic diversity and changes in the 56 large urban areas from 1980 to 2020. It begins with examination of summary statistics including averages and ranges of values over time. Next come lists of the urban areas with the highest and lowest levels of exposure and evenness diversity, contrasting these measures at the extremes and raising the issue of variation by region. Simple exploratory models examine the relationships of several measures of urban area population and diversity to levels and changes in exposure and evenness diversity to conclude the section.

Table 1 presents summary statistics for the exposure and evenness diversity indexes for the urban areas from 1980 to 2020. The means for both indexes increase significantly over the 40-year period, from 28 to 57 for exposure diversity and 58 to 78 for evenness diversity. The means for evenness diversity are higher than the means for exposure diversity, which must be the case as the exposure diversity is divided by the larger urban area diversity to give evenness diversity. These differences are larger in the early years, a 30 point difference in the means in 1980 compared to a 20 point difference in 2020.

The results in Table 1 show very large variations across the urban areas. For example, in 1980 the urban areas range from 12 to 59 for the exposure diversity index and from 35 to 96 for the evenness diversity index. Comparing the ranges and seeing how they change over time is made easier looking at the information plotted in Figure 1. The minimum, median, and maximum values are plotted by year for both exposure

Table 1. Summary statistics for exposure and evenness diversity, 1980-2020

Year	Mean	Minimum	First quartile	Median	Third quartile	Maximum
Exposure diversity						
1980	28.1	11.6	18.8	25.4	35.2	58.5
1990	34.7	16.1	23.5	33.3	44.6	63.4
2000	42.8	20.2	33.0	42.0	51.7	69.1
2010	50.9	25.6	42.2	51.9	58.9	74.0
2020	57.2	32.8	49.7	58.8	64.7	80.1
Evenness diversity						
1980	58.2	34.5	46.4	55.6	69.0	96.0
1990	64.8	33.8	54.7	65.5	77.7	95.3
2000	69.4	41.2	61.8	71.1	77.7	92.0
2010	74.1	53.1	69.0	75.5	80.8	92.6
2020	77.8	60.4	74.1	79.1	83.6	92.7

diversity and evenness diversity. The former are in red, the latter in blue. Areas between the minima and maxima are likewise shaded in those colors, with the area of overlap darker.

Some observations: The exposure index varies across a lower range of values than the evenness index, of course, but there is substantial overlap. Not only do the average values increase over time but the minima and maxima do as well. For exposure diversity, the range is fairly steady with some increase below the median and some decrease above the median over time. For evenness diversity, the range gets smaller, especially above the median with the large difference between the median and maximum in 1980 shrinking to a much smaller difference by 2020 (with the maximum constrained by the largest possible value of 100).

A fascinating coincidence occurs with the extreme values for one index closely coinciding with the medians for the other index. The minimum value for evenness diversity is nearly the same as the median exposure diversity except in 1980. Likewise, the maximum exposure diversity coincides with the median evenness diversity. Nothing in the way the indexes are calculated creates this relationship.

The wide range of diversity index values suggests insight may be gained by looking at the urban areas at the extremes. Table 2 lists the six urban areas (about ten percent of all areas) with the highest exposure diversity in 2020 and the six areas with

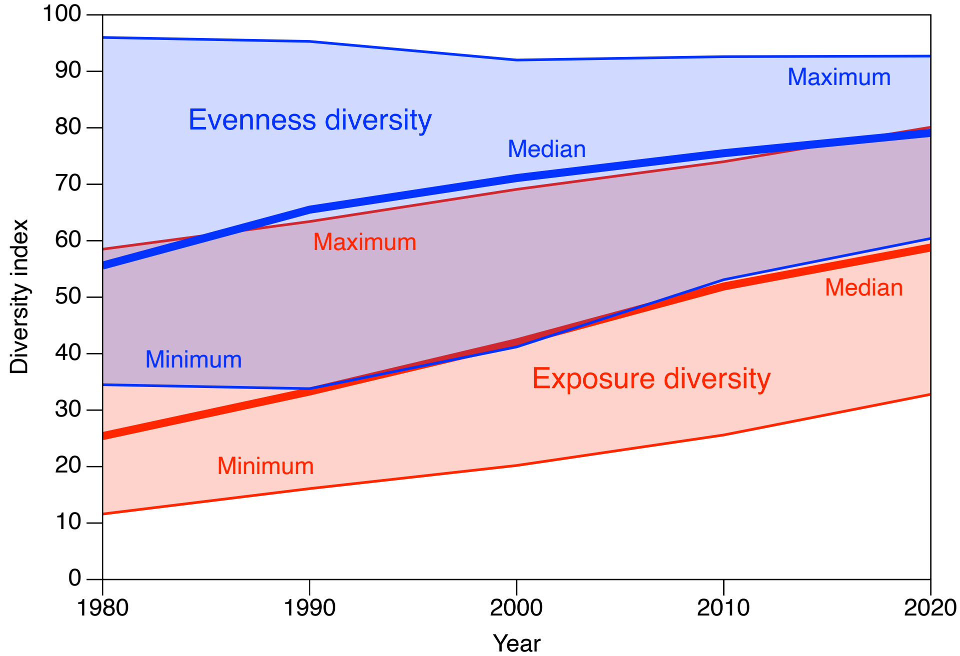

Figure 1. Ranges of exposure and evenness diversity, 1980-2020

the lowest. The evenness diversity and the overall urban area diversity are also listed for comparison. In the same way Table 3 lists the areas with the highest and lowest evenness diversity.

Starting with the areas having the highest exposure diversity in Table 2, all are rapidly growing areas located in the sunbelt. The three most diverse areas are in the West, the next two are in Texas, with Orlando the sixth area. The areas are in the top half of the 56 urban areas in terms of population, with three among the ten largest areas. Overall urban area diversity is also high, with five of the six areas in the top ten on that measure. This is not surprising, given the high correlation of 0.85 between exposure diversity and urban area diversity. But for evenness diversity, while all of the areas are above average, they are not among the highest on this measure. Evenness diversity is necessarily related to exposure diversity, but this relationship is much weaker with a correlation of 0.36.

The urban areas with the lowest levels of exposure diversity could not be more different. Five of the six are older areas in the rustbelt, in the Northeast and Midwest. The one exception is El Paso with the lowest exposure diversity. El Paso is an outlier with 85 percent of the population Latino in 2020, producing the lowest levels of both urban area diversity and exposure diversity. The areas are also below average on evenness diversity with the exception, once again, of El Paso which has extremely high

Table 2. Urban areas with the highest and lowest exposure diversity, 2020

Urban area	*Exposure diversity*	*Evenness diversity*	*Urban area diversity*
Las Vegas	80.1	87.6	91.5
Sacramento	76.5	86.8	88.2
San Francisco-Oakland-San Jose	75.2	81.0	92.9
Houston	72.3	78.1	92.6
Dallas-Fort Worth	72.1	78.4	92.0
Orlando	72.0	83.4	86.4
...	...	...	...
Cincinnati	41.1	74.3	55.3
Buffalo	40.8	69.0	59.1
St Louis	40.1	62.7	63.8
Cleveland-Akron	38.6	64.0	60.3
Pittsburgh	32.8	77.0	42.6
El Paso	32.8	90.9	36.0

evenness diversity. But the area's low exposure diversity is high relative to the very low overall urban area diversity.

The list of areas with the highest evenness diversity shown in Table 3 presents a very different picture. It includes both the area with the highest exposure diversity, Las Vegas, and the area with the lowest exposure diversity, El Paso. For Las Vegas, its exposure diversity is so high that even adjusting for the urban area diversity, it remains fifth-highest in terms of evenness diversity. El Paso is just the opposite. While the exposure diversity is extremely low, it is not that much lower than the also extremely low urban area diversity, boosting the evenness diversity to second-highest. The other four areas with the highest evenness diversity have levels of exposure diversity that are more in the middle of the distribution, neither very high nor very low. All but El Paso are in the West and El Paso is nearly as far west as Albuquerque and is farther west than Denver.

The list of areas with low evenness diversity share more with the areas having low exposure diversity. Cleveland-Akron and St Louis are among the lowest areas on both lists. While the other areas with low evenness diversity are different from those with low exposure diversity, they are similar types of areas. Of those areas with the lowest exposure diversity, all but El Paso are older urban areas in the Northeast and

Table 3. Urban areas with the highest and lowest evenness diversity, 2020

Urban area	*Evenness diversity*	*Exposure diversity*	*Urban area diversity*
Portland	92.7	56.2	60.6
El Paso	90.9	32.8	36.0
Seattle-Tacoma	90.3	67.5	74.7
Salt Lake City-Provo	88.7	47.5	53.5
Las Vegas	87.6	80.1	91.5
Albuquerque	87.0	63.8	73.3
...	...	...	...
Chicago	64.4	56.2	87.2
Cleveland-Akron	64.0	38.6	60.3
St Louis	62.7	40.1	63.8
Milwaukee	62.3	47.0	75.4
Birmingham	61.5	47.0	76.5
Detroit	60.4	42.4	70.2

Midwest. Five of the six areas with the lowest evenness diversity are older areas in the Midwest. All of the lowest evenness diversity areas have exposure diversity below the mean, with all but Chicago in the bottom quartile. Urban area diversity is near the mean or below with again the exception of Chicago, which is more diverse.[2]

The lists of areas with the highest and lowest levels of exposure and evenness diversity clearly suggest patterns of variation by region. Table 4 shows the mean values of those measures in 2020 by region and the mean changes from 1980 to 2020. The four census regions are used with one exception. The census puts the Washington-Baltimore area in the South. As it is part of the huge, nearly continuous area of urban development extending to north of Boston, the area is considered to be in the Northeast in the current research.

[2] A study (Frey 2022) lists the ten major metropolitan areas with the highest levels of black-white segregation in 2020 using the index of dissimilarity. Five of the six areas with the lowest levels of evenness diversity were on the list. El Paso was not included among the major metropolitan areas included. Among the areas with. the highest evenness diversity, Seattle, Las Vegas, Portland, and Salt Lake City appeared on at least one of the lists of ten areas with the lowest black-white, Latino-white, and Asian-white segregation, with the first two areas on two of the lists. Like El Paso, Albuquerque was not among the major metropolitan areas in the study.

Table 4. Mean exposure and evenness diversity in 2020 and change from 1980 to 2020 by region

	Northeast	*Midwest*	*South*	*West*
Exposure diversity				
Diversity 2020***	52.0	48.0	59.5	64.1
Change 1980-2020	29.9	29.2	30.6	26.0
Evenness diversity				
Diversity 2020***	73.0	71.7	78.9	84.0
Change 1980-2020***	18.6	22.8	26.4	5.0

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

Average exposure diversity is highest in the West at 64 and lowest in the Midwest at 48. The South is not as high and the Northeast not as low as the regions at the extremes. The differences are significant at the 0.001 level. However the change in exposure diversity from 1980 to 2020 does not vary significantly by region.

A similar pattern holds for evenness diversity even though urban areas can have widely varying evenness and exposure diversity. Again, the West is on average most diverse at 84 and the Midwest least at 72, with the other two regions in the same relative positions. And these difference are highly significant. But unlike change in exposure diversity, the mean increases in evenness diversity vary greatly across the regions. Average diversity in the West increased by only five points, very small compared to the other regions. This despite the West having the highest evenness diversity in 2020. The West starts the period high and remains high. In contrast, the other three regions experience increases ranging from 19 to 26 points with the South highest. Differences in the change in evenness diversity are also highly statistically significant.

Regression models are used to examine the relationships of some basic characteristics of the urban areas to the measures of diversity. The urban area characteristics are the log of population and its change, the percent of the population in each of the racial and ethnic groups. and the changes in those percentages. These are used to predict exposure diversity in 2020 and its change from 1980 to 2020 and evenness diversity and its change. Urban area variables for 2020 are used in the models predicting diversities in the year and the variables for 1980 at the start of the period are used for predicting the changes from 1980 to 2020.

The percentages in the racial and ethnic groups sum to 100 percent. They are perfectly collinear and a model including all four variables cannot be estimated. This must be addressed by not including one or more of the variables in the model. Given that whites are the largest group in the majority of the areas, examining the relationship of diversity to percent white and its change is desirable. But the high percent white in some areas creates a problem of multicollinearity with high variance inflation factors with even one other group included in the model. It was also desirable to see the relationships of diversity to the percentages and changes in the other groups. To accomplish this, two models are estimated for each measure of diversity and change. The first includes percent white and its change. The second includes the percentages in the other three groups and their changes. Note that both models are essentially being estimated using percent white and percent nonwhite, with the latter broken down into the percentages for the other groups, providing more detailed information and producing better predictions.

The first two columns in Table 5 give the results for the models predicting exposure diversity in 2020. The change in the logarithm of population from 1980 to 2020 has a positive and highly significant effect. This makes sense, as greater population change increases the possibilities for more diversity. Among the coefficients for the percentages in the racial and ethnic groups in 2020, percent white is negative and significant. Fewer whites generally means greater diversity. The coefficients for the

Table 5. Models predicting exposure diversity 2020 and change in exposure diversity, 1980-2020

Independent variable Urban area characteristic	*Exposure diversity 2020*		*Exposure diversity 2020*		*Change in exposure diversity 1980-2020*		*Change in exposure diversity 1980-2020*	
Log population 2020/1980	1.377		-1.471		1.812		0.921	
Change log population 1980-2020	9.341	***	8.197	**	9.904	**	9.657	***
Percent white 2020/1980	-0.171	*	—		0.512	***	—	
Percent black 2020/1980	—		0.346	***	—		-0.074	
Percent Latino 2020/1980	—		0.059		—		-0.677	***
Percent Asian 2020/1980	—		0.222	*	—		-0.457	***
Change percent white 1980-2020	-0.441	*	—		-0.115		—	
Change percent black 1980-2020	—		-0.067		—		0.621	**
Change percent Latino 1980-2020	—		0.609	*	—		0.314	*
Change percent Asian 1980-2020	—		1.516	***	—		1.025	***
Constant	31.820		48.559	*	-42.955	*	9.583	
R^2	0.554	***	0.725	***	0.603	***	0.855	***

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

other three groups are all positive, with percent black and percent Asian significant. Greater proportions in these groups are associated with more diversity. And then for the changes in percent in these groups from 1980 to 2020, the relationships are similar. Greater decreases in whites are significantly associated with higher levels of diversity. Increases in Latinos and Asians were likewise significantly associated with greater diversity. The coefficient for change in percent black was negative but very small and not significant, and blacks experienced by far the smallest change in their share of the population over the forty-year period.

Continuing with the prediction of the change in exposure diversity from 1980 to 2020, the results are in the final columns of Table 5. The change in the log of population is again positively associated with the change in diversity, with the rationale the same as for the level of diversity. But for the change in diversity, the initial percentage of the population white in 1980 is positively related to diversity change. It is likely the case that areas with greater proportions white had more opportunity for the decline in the percent white, leading to increased diversity. The negative coefficient for change in percent white, while not significant, is consistent with this. For the other three groups,

Table 6. Models predicting evenness diversity 2020 and change in evenness diversity, 1980-2020

Independent variable *Urban area characteristic*	*Evenness diversity 2020*		*Evenness diversity 2020*		*Change in evenness diversity 1980-2020*		*Change in evenness diversity 1980-2020*	
Log population 2020/1980	-2.993	*	-3.565	***	1.304		1.421	
Change log population 1980-2020	8.697	***	9.639	***	8.388	*	6.631	
Percent white 2020/1980	0.030		—		-0.022		—	
Percent black 2020/1980	—		-0.357	***	—		0.516	***
Percent Latino 2020/1980	—		0.097		—		-0.057	
Percent Asian 2020/1980	—		0.004		—		-0.257	*
Change percent white 1980-2020	-0.227		—		0.752	**	—	
Change percent black 1980-2020	—		0.157		—		-0.379	
Change percent Latino 1980-2020	—		-0.264		—		-0.304	
Change percent Asian 1980-2020	—		0.566	*	—		-0.784	*
Constant	109.408	***	127.178	***	11.954		-1.603	
R^2	0.396	***	0.709	***	0.224	*	0.625	***

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

the opposite pattern holds and can be explained in the same way. The coefficients for percent black, Latino, and Asian are all negative with the latter two highly significant. Larger starting shares in these groups provided less opportunity for increases and therefore greater diversity. Increases in the proportions in percentages over time were not surprisingly associated with greater increases in diversity.

Table 6 provides analogous results for evenness diversity in 2020 and the change in evenness diversity from 1980 to 2020. As before, the change in the log of population is significantly related to evenness diversity. But the percent white in 2020 and the change in the percent white are not significant, in contrast to exposure diversity. The only percentage of any group significantly related to evenness diversity is the percent black, and the sign is negative, just the opposite of what was observed for exposure diversity. For the change in percent variables, only the change in percent Asian was significant and this was also positive, which is the one case consistent with the models for exposure diversity. The negative sign and the lack of significance for the change in percent Latino was especially surprising, as this is the group experiencing the largest change in percent over the period.

For the change in evenness diversity, log of population change is again positively related, though the levels of significance are lower than in all of the other models, with one coefficient not significant. Percent white in 1980 is negative and not significant for evenness diversity while it is highly significant and positively related to change in exposure diversity. Greater change in percent white is significantly associated with larger increases in evenness diversity which is a curious result to say the least. Coefficients for the changes in percent for the other three groups are all negative, compared to uniformly positive and significant coefficients for exposure diversity. Why decreases in the percentages in these groups should be associated with increases in evenness diversity is not clear.

Exposure diversity in the urban core and suburban periphery

A longstanding, simple approach for examining the residential distribution of members of racial and ethnic groups is the comparisons of their percentages of the populations in the central cities and suburbs. As with research on residential segregation, this begins with a focus on percentages black and white, with a paper by Sharp and Schnore (1962) an early example. More recently, in the book *Diversity Explosion*, Frey (2015) examines the percentage distributions of four racial and ethnic groups at various geographic levels including central cities and suburbs.

The idea underlying the central city-suburb division is that the central cities are the older parts of the urban area while the suburbs are newer, having been developed more recently. The problem is that central cities include widely varying portions of their urban areas and their boundaries can vary over time, as have the very definitions of central cities. An alternative approach is taken in the urban patterns research. The urban core is defined as the extent of the urban area in 1940, a fixed area that is consistently defined across urban areas. The suburban periphery is that territory developed and added to an urban area since that time. In other words, the suburbs are the newer areas primarily built since World War II.

An earlier paper (Ottensmann 2023c) compares overall levels of racial and ethnic diversity over time in the urban core and the suburban periphery. Average diversity starts out much lower in 1980 in the suburbs than in the core. While both increase over time, diversity in the suburbs increases more rapidly, narrowing the gap, with average diversity in the periphery in 2020 nearly has high as diversity in the core. These are the overall levels of diversity in these areas, not diversity within core and suburban neighborhoods (census tracts). This section focuses on the exposure and evenness diversity in the urban core and suburban periphery.

Table 7 lists the mean exposure and evenness diversity in the urban core and suburban periphery by year. For reference, it also lists the overall levels of diversity for these areas. Mean exposure diversity, starts out low in 1980, especially in the suburban periphery, 25 in the suburbs and 34 in the urban core. The means increase substantially over the forty-year period becoming essentially identical, with values of 56 and 57 in 2020.

Further comparisons of exposure diversity with evenness diversity are easier to see in Figure 2, where the means from Table 1 are plotted over time. Focus first on the purple lines, for the measures of diversity in the urban core, and the orange lines, for diversity in the suburban periphery. The two lower lines show the mean core and suburban exposure diversity. As discussed, the suburbs start out lower and the lines increase and converge by 2010 and 2020.

Table 7. Mean exposure and evenness diversity and overall urban area diversity in the urban core and suburban periphery, 1980-2020

Year	Exposure diversity		Evenness diversity		Overall area diversity	
	Urban core	Suburban periphery	Urban core	Suburban periphery	Urban core	Suburban periphery
1980	33.5	25.3	56.3	72.6	61.3	36.3
1990	39.3	32.6	60.1	77.1	66.2	43.2
2000	44.8	41.5	64.0	78.6	70.7	53.7
2010	50.2	50.2	68.7	80.4	73.8	63.2
2020	56.1	56.7	73.9	82.0	76.6	69.6

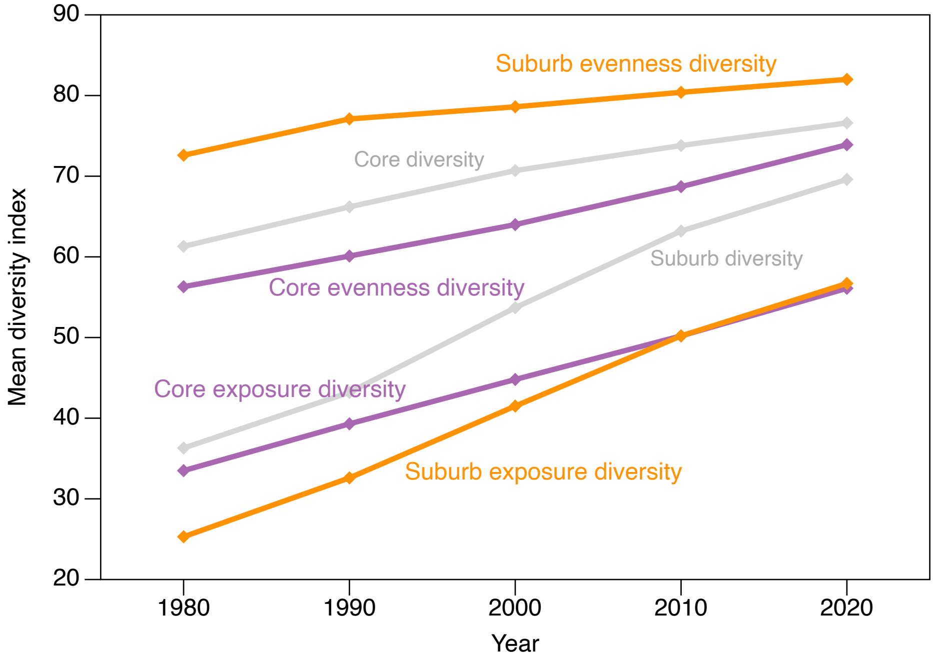

Figure 2. Mean exposure and evenness diversity and overall urban area diversity in the urban core and suburban periphery, 1980-2020

Evenness diversity is higher, as it must be. Core evenness diversity increases very much in tandem with exposure diversity, with the difference decreasing somewhat, from 23 to 18. Evenness diversity in the suburban periphery is very different. It is very high, between 73 and 82, much higher than evenness diversity in the core in every year, especially in the early years. The values in the early years are especially striking. In 1980, exposure diversity in the suburbs was only 25 while evenness diversity was 73. The two measures are providing radically different pictures of racial and ethnic diversity in the suburbs.

To understand what is causing this, turn now to the overall diversity in the core and suburbs shown by the gray lines in Figure 2. Overall diversity is the maximum possible value for exposure diversity, and evenness diversity is exposure diversity divided by overall diversity. Mean overall diversity was extremely low in the suburbs in 1980, only 36, compared to 61 in the urban core. So while the exposure diversities in the suburbs in 1980 are very low, they are being divided by the very low overall diversities in the suburbs, producing the high mean evenness diversity of 74. Over time, mean overall diversity in the suburbs nearly doubled, to 70 while exposure diversity more than doubled, from 25 to 57. The net result was the modest increase in mean evenness diversity in the suburbs over time.

The discussion of exposure and evenness diversity in the core and suburbs has focused on the mean values over time. A great deal of variation exists among the large urban areas. To consider this, Table 8 presents more complete summary statistics for those measures of diversity in 2020. The ranges are very large. Exposure diversity in the urban core ranges from 20 to 90, with the range in the suburbs only somewhat smaller. The ranges for evenness diversity are necessarily smaller as the minimum must exceed exposure diversity and the maximum is bounded by 100. Even so, evenness diversity in the core ranges from 43 to 97, with much less variation in the suburban periphery. Values for the first and third quartiles show, however, that half of the areas have diversity values clustered in much smaller ranges, from 51 to 62 for exposure diversity in the core, for example.

Table 8. Summary statistics for exposure and evenness diversity in the urban core and suburban periphery, 2020

	Mean	*Minimum*	*First quartile*	*Median*	*Third quartile*	*Maximum*
Exposure diversity						
Core diversity 2020	56.1	20.4	51.0	57.1	62.2	89.9
Suburb diversity 2020	56.7	25.5	50.2	59.0	66.6	80.1
Evenness diversity						
Core diversity 2020	73.9	43.4	67.9	73.2	80.5	96.8
Suburb diversity 2020	82.0	65.6	78.9	82.9	86.9	95.1

Tables listing the urban areas with high and low levels of diversity in the urban core and suburban periphery are not included. But the areas at the extremes, some of which are clear outliers, do give further insight into diversity in the core and suburbs. The urban areas with the highest exposure diversity in the core are Las Vegas at 90 and Sacramento and San Francisco-Oakland-San Jose at 77. Below that is another gap down to numbers of areas with exposure diversity at 66 and 67. At the bottom are El Paso at 20 and Detroit and San Antonio at 34 and 36.

Shifting to evenness diversity in the core, Las Vegas again tops the list at 97 followed by Portland at 95. While Las Vegas also has high exposure diversity, Portland's exposure diversity was only 51, right at the first quartile. Third, at 94, is El Paso, once again at the bottom in terms of exposure diversity and near the top in terms of evenness diversity. The areas with the lowest evenness diversity in the core are Detroit, Milwaukee, and Chicago, with values ranging from 43 to 56. Exposure diversity is well below average for the latter two, with Detroit the second lowest behind El Paso.

The three areas with the highest exposure diversity in the suburban periphery are the same as for the core, Las Vegas, Sacramento, and San Francisco-Oakland-San Jose. The urban areas with the lowest exposure diversity in the suburbs are led by Pittsburgh and Buffalo at 25 and 29. Six of the seven least diverse areas are in the rustbelt. The other is not surprisingly El Paso.

Evenness diversity in the suburbs presents a very different picture compared to exposure diversity in the suburbs. Rochester and Portland are the most diverse at 95 and 92, with the former having very low exposure diversity at 38 while Portland is just above the mean at 58. El Paso is again high, in third. Pittsburgh and Buffalo, the least diverse areas in terms of exposure diversity, are in eighth and ninth places on the list for evenness diversity. The areas with the lowest evenness diversity are even more of a mix. St. Louis also has low exposure diversity, but Miami-Fort Lauderdale-West Palm Beach and Washington-Baltimore are among the areas with the highest exposure diversity.

Average levels of diversity for the entire urban areas vary significantly by region. Examination of the areas placing high and low for the core and suburbs suggests such variation might exist for diversity in those areas as well. Table 9 gives the means by region for exposure and evenness diversity for the urban core and suburban periphery. Average levels of exposure diversity in the core are highest in the West, followed by the Northeast. Diversity in the Midwest and South are lower. Differences are statistically significant. For the suburbs, difference in exposure diversity are highly significant but they follow a different pattern. Once again mean diversity is highest in the West but now it is the South that is close. Average exposure diversity in the Northeast and Midwest are much lower.

Table 9. Mean exposure and evenness diversity for the urban core and suburban periphery by region, 2020

	Northeast	*Midwest*	*South*	*West*
Exposure diversity				
Urban core*	59.5	52.1	53.0	63.1
Suburban periphery***	46.0	46.4	60.5	64.8
Evenness diversity				
Urban core***	71.1	65.5	72.7	85.0
Suburban periphery	84.5	80.2	80.8	84.3

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

Mean evenness diversity in the urban core is highest by far in the West. Diversity in the Northeast and South are similar and substantially lower, with the mean level in the Midwest being much, much lower, 66 versus 85. These differences are significant. Breaking the pattern, average levels of evenness diversity do not vary that much across the regions and the difference are not statistically significant.

In the prior study, mean levels of overall diversity in the suburban periphery were much lower in 1980 than levels in the urban core. Diversity increased over time and by 2020, diversity in the suburbs approached diversity in the core. To focus on these differences, the ratio of suburban diversity to core diversity was calculated for each area and the means summarize these results for each year. These are shown in the final column of Table 10. The mean suburb-core diversity ratio starts out at 0.59 in 1980 and climbs to 0.93 in 2020 as diversity in the suburbs approaches diversity in the core.

Table 10. Mean suburban periphery-urban core ratios for exposure and evenness diversity and mean ratios for overall area diversity, 1980-2020

Year	*Mean suburban periphery-urban core ratio*		
	Exposure diversity	*Evenness diversity*	*Overall area diversity*
1980	0.75	1.37	0.59
1990	0.85	1.35	0.67
2000	0.97	1.28	0.79
2010	1.04	1.20	0.89
2020	1.04	1.13	0.93

A similar comparison is made of the suburb-core ratios for exposure diversity and evenness diversity, with the means for those ratios over time also shown in Table 10. The mean suburb-core exposure diversity ratios are somewhat higher than the mean ratios for overall area diversity, but they follow a similar pattern. The mean ratio starts out lower in 1980, with the average suburb less diverse than the core. The ratios then increase so that by the end of the period, the mean suburb-core ratios exceeded one, with the suburb having higher exposure diversity than the core.

The mean ratios of evenness diversity in the suburban periphery to evenness diversity in the urban core could not have a more different pattern. The mean ratio starts out high in 1980 at 1.37. Evenness diversity is higher in the suburbs than in the core. A major factor contributing to this is the much lower overall diversity in the suburbs, causing the evenness diversity to be higher. The mean suburb-core evenness diversity ratio declines steadily over time to 1.13 in 2020. Suburban evenness diversity still tends to be higher than evenness diversity in the core, but the differences are far less.

Conclusions

A major result is the consistent increase in average levels of racial and ethnic residential diversity across the large urban areas over time. This holds for both exposure diversity and evenness diversity and for the entire urban areas and the urban cores and suburban peripheries. These trends corresponds to the increases in diversity at the national, state, and urban area level and confirms that the increases extend down to the neighborhood level.

This research employed two measures of diversity at the neighborhood level. For exposure diversity, a measure of racial and ethnic diversity in each census tract is averaged to produce an overall measure of diversity within neighborhoods for the urban area and for the urban core and suburban periphery. The highest possible value for exposure diversity is the overall level of diversity in the larger area. Evenness diversity is then exposure diversity divided by that overall level of diversity, diversity at the neighborhood level relative to diversity in the larger area. Evenness diversity has a maximum possible value of one when the distribution across the racial and ethnic groups in each census tract is the same as the distribution in the larger area.

Exposure and evenness diversity are distinct measures that provide different perspectives on levels of diversity within neighborhoods in the urban areas. The correlation between exposure and evenness diversity is 0.36, meaning that only 13 percent of the variation in evenness diversity is accounted for by the variation in exposure diversity.

The differences are obvious in some of the results presented here. The lists of the areas with the highest and lowest exposure and evenness diversity in Tables 2 and 3 show significant inconsistencies. Most dramatic is El Paso, which has the lowest level of exposure diversity and the second highest level of evenness diversity. The differences are especially clear among the areas with the highest evenness diversity. The areas with the highest evenness diversity include areas that have levels of exposure diversity that are very high, very low, and in the middle. But areas with high levels of exposure diversity cannot have extremely low levels of evenness diversity. Similar differences exist for the lists of highest and lowest exposure and evenness diversity in the urban core and suburban periphery, not included in this paper.

Differences arise again with the models predicting exposure and evenness diversity in 2020 and changes from 1980 to 2020 using urban area populations, percent in the racial and ethnic groups, and the changes in those variables, with the results in

Tables 5 and 6. For exposure diversity the results seem reasonable and consistent with expectations. Lower percent white and higher percent in the other three groups are associated with higher levels of exposure diversity. Decreases in percent white since 1980 and increases in percent Latino and Asian over the period are associated with higher levels of exposure diversity. The models predicting change in exposure diversity over the period are likewise reasonable.

In the models predicting evenness diversity in 2020, percent white and the change in the percent white are not statistically significant. Only two variables are statistically significant for the other three groups, percent black and and change in percent Asian. Percent black is negatively related to evenness diversity, just the opposite from exposure diversity and difficult to explain. The only statistically significant and consistent predictor is change in percent Asian, which is associated with higher levels of evenness diversity.

A third example of differences between exposure and evenness diversity involves the mean levels over time for diversity in the suburb periphery, in Table 7 and Figure 2. Exposure and evenness diversity in the urban core is not an issue. Evenness diversity is of course higher than exposure diversity, as it must be. Both increase in similar ways. Mean suburban exposure diversity in 1980 was extremely low, only 25, lower than core diversity. It more than doubled over the period. However mean evenness diversity in the suburbs in 1980 is a very high 73, much higher than for the core, and it increases only modestly to 82 by 2020. The very low levels of overall diversity in the suburbs cause the evenness diversity to be this high in the earlier years, producing this result.

This highlights what is driving the differences between exposure and evenness diversity. Evenness diversity is exposure diversity divided by the overall diversity of the larger area. Urban areas with low levels of overall diversity will thus have their evenness diversity increased more than other areas. This produces a curious situation. Duncan and Duncan (1955) favor segregation indexes that are independent of the racial distribution in the urban area, measures which Massey and Denton (1988) say reflects "evenness" in the distributions of the groups across the tracts. The evenness diversity measure used here is one examples of such a measure. And it is indeed essentially independent of urban area diversity, with a correlation of only 0.18, not statistically significant. However the value of the evenness diversity index is computed using urban area diversity. In that sense evenness diversity does depend on urban area diversity.

Theil's information theory index is similar to the evenness diversity index, using the entropy index for the measure of tract diversity and subtracting the area diversity from one to produce a measure of segregation. The index is initially used to study the levels of racial segregation across schools in a school system (Theil and Finezza 1971; Theil 1972). Such a measure of relative segregation is very appropriate in this context. The issue is the extent to which the school system has pursued policies that may have made the schools more segregated and less diverse than they could have been. An even distribution of the students in the racial groups across the schools would constitute maximum evenness diversity and zero segregation. That outcome would be determined by the overall racial distribution of all of the students in the schools, which is assumed to be a given over which the school system has no control.

This sort of argument does not transfer quite so neatly to residential segregation (or diversity) by race and ethnicity as would be measured by Theil's information theory index of segregation or by the evenness diversity index used in this paper. An evenness segregation index might be seen as the extent to which the housing market is causing the distribution of the population by race to be more segregated than it had to be, less diverse than the maximum diversity possible. This is certainly one conception of segregation. But can one reasonably compare the operation of a housing market in an area with a highly diverse population with the market in an area in which eighty percent of the population are members of a single group?

In some of the results, the differences between exposure and evenness diversity are most apparent among urban areas having lower levels of diversity. This suggests considering the implications for the most extreme example. In 1980, Minneapolis-St.Paul was 95 percent white, highest among all of the urban areas. The level of urban area diversity was 14. If every tract had exactly the same racial and ethnic distribution as the urban area, exposure diversity would have the maximum value of urban area diversity, also 14. This would produce a value for evenness diversity of 100 and therefore zero evenness segregation. Is it reasonable to say that a tract that is 95 percent white is completely integrated?

This is not to argue that evenness diversity and segregation are not useful measures. Residents of an urban area will evaluate levels of diversity and segregation in a neighborhood in different ways. One person could base the judgement on the mix of racial and ethnic groups within the neighborhood and conclude that a neighborhood that is 80 percent white is not highly diverse. Another person, a resident of the Pittsburgh area in 2020, could look at the same neighborhood and conclude that since it has nearly the same composition as the urban area as a whole, which was 81 percent white, this is a diverse neighborhood.

References

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.

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.

Duncan, Otis D., and Beverly Duncan. 1955. A methodological analysis of segregation indices. *American Sociological Review* 20: 210-217.

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

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

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

Frey, William H. 2015. *Diversity Exposition: How New Racial Demographics Are Remaking America*. Brookings Institution Press.

________. 2022. A Census portrait of America's largest metro areas: population growth, diversity, segregation and youth. Brookings Mountain West. Downloaded from *https://digitalscholarship.unlv.edu/brookings_policybriefs_reports/11/* on November 18, 2023.

________, and Reynolds Farley. 1996. Latino, Asian, and black segregation in U.S. metropolitan areas: are multi-ethnic metros different? *Demography* 33: 35-50.

Galster, George. 1998. A stock/flow model of defining racially integrated neighborhoods. *Journal of Urban Affairs* 20: 43-51.

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.

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., and Nancy A. Denton. 1987. Trends in the residential segregation of blacks, Hispanics, and Asians. *American Sociological Review* 52: 802-825.

________. 1988. The dimensions of residential segregation. *Social Forces* 67: 281-315.

________. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press.

Massey, Douglas S., and Jonathan Tannen. 2018. Suburbanization and segregation in the United States: 1970-2010. *Ethnic and Racial Studies* 41: 1594-1611.

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*.

________. 2023b. Density and sprawl in the older and newer parts of large urban areas in the U.S., 1970-2020. Available at *https://urbanpatternsblog.files.wordpress.com/2023/06/density-and-sprawl-in-older-and-newer.pdf*.

________. 2023c. Racial and ethnic diversity in large urban areas in the U.S., 1980-2020. Available at *https://urbanpatternsblog.files.wordpress.com/2023/11/racial-and-ethnic-diversity.pdf*.

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.

Reardon, Sean F., John T. Tun, and Tamela Mcnulty Eitle. 2000. The changing structure of school segregation: measurement and evidence of multiracial metropolitan-area school segregation, 1989-1995. *Demography* 37: 352-364.

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

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

Sharp, Harry, and Leo F. Schnore. 1962. The changing color composition of metropolitan areas. *Land Economics* 38: 169-185.

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: 1-25.

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

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.