**Structural versus Ethnic Dimensions of Housing Segregation**

**by**

**Richard Sander**
**University of California, Los Angeles**

**Yana Kucheva**
**CUNY**

**CES 16-22 March, 2016**

The research program of the Center for Economic Studies (CES) produces a wide range of economic analyses to improve the statistical programs of the U.S. Census Bureau. Many of these analyses take the form of CES research papers. The papers have not undergone the review accorded Census Bureau publications and no endorsement should be inferred. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Republication in whole or part must be cleared with the authors.

To obtain information about the series, see www.census.gov/ces or contact Fariha Kamal, Editor, Discussion Papers, U.S. Census Bureau, Center for Economic Studies 2K132B, 4600 Silver Hill Road, Washington, DC 20233, CES.Papers.List@census.gov. To subscribe to the series, please click here.

**Abstract**

Racial residential segregation is still very high in many American cities. Some portion of segregation is attributable to socioeconomic differences across racial lines; some portion is caused by purely racial factors, such as preferences about the racial composition of one's neighborhood or discrimination in the housing market. Social scientists have had great difficulty disaggregating segregation into a portion that can be explained by interracial differences in socioeconomic characteristics (what we call structural factors) versus a portion attributable to racial and ethnic factors. What would such a measure look like? In this paper, we draw on a new source of data to develop an innovative structural segregation measure that shows the amount of segregation that would remain if we could assign households to housing units based only on non-racial socioeconomic characteristics. This inquiry provides vital building blocks for the broader enterprise of understanding and remedying housing segregation.

---

Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

## Introduction

Urban housing segregation, particularly among African-Americans, is still very high in many major American cities by any measure. In the sixteen metropolitan areas with the largest black populations in the United States, roughly 40% of African-Americans live in essentially all-black neighborhoods, and fewer than a quarter live in neighborhoods that are mostly white or Asian. Since every year brings new evidence that levels of housing segregation predict a wide variety of bad outcomes, understanding and addressing segregation should be a major concern for public policy.

One of the half-dozen or so biggest questions in the study of racial segregation concerns the degree to which it is fundamentally structural or fundamentally ethnic (Massey and Denton 1993; Massey and Fischer 2000; Wilson 1996; Quillian 2002). By "structural" we mean driven by objective characteristics in households of different races. For example, black households in nearly all metropolitan areas have substantially lower mean and median incomes than white households. Since neighborhoods are far more homogeneous in terms of income than are entire metropolitan areas, the black/white differences in incomes will automatically produce some geographic stratification and thus housing segregation. Black and white (but also Hispanic) households have differing distributions not only of income, but in the age of the householder, the size of the household, the presence of children – and each of these has an influence on housing choice and availability. To the extent that individual neighborhoods encompass only a limited range of housing choices, these demographically distinct households will live apart, and some degree of racial segregation will logically follow.

Thus, ever since federal and state governments widely endorsed principles and policies against racial discrimination in housing in the late 1960s, a major battleground in fair housing

has been the question of "inclusionary" zoning: how to introduce more rental housing and/or low-and-moderate income housing into suburbs and neighborhoods that are predominantly made up of affluent families in detached, single-family homes. In many major court cases, fair housing advocates have painted racial segregation as nearly synonymous with structural factors that limit access of the disadvantaged to affluent neighborhoods.[1]

Racial housing segregation can and has of course been caused by purely "racial" factors as well. Early 20th century efforts by Southern and Border cities to create racial zoning are the most obvious example.[2] The racially restrictive covenants and very high levels of racial discrimination that were prevalent in mid-century northern cities are others. Even in the absence of deliberate attempts to restrict housing choices based on race, widespread preferences about the racial composition of one's neighborhood can produce patterns of avoidance that lead to high levels of racial segregation (Schelling 1971; Muth 1986).

Social scientists have developed many ways of measuring levels of housing segregation. For example, the index of dissimilarity between two groups (say blacks and whites) in a city measures the proportion of whites who would need to move into a different neighborhood to achieve the same racial balance in all neighborhoods. An index measure of 100% (or 1.0) indicates complete segregation; a measure of 0 indicates complete integration. The index of

---

[1] See, for example, Village of Arlington Heights v. Metropolitan Housing Development Corp., 429 U.S. 252 (1977); Huntington Branch NAACP v. Town of Huntington, 844 F.2d 926 (2nd Cir 1988); South Burlington County NAACP v. Mt. Laurel, 336 A. 2d 713 (1975).

[2]Racial zoning was declared unconstitutional by the Supreme Court in 1917, but attempts to use such tools nonetheless persisted in some southern cities until the middle of the 20th century Connerly (2005).

black/white dissimilarity, averaged across major urban areas with large black populations, was .84 in 1970 and fell to .69 by 2010. Despite this decline, large income and wealth disparities between blacks and whites persist, and the degree of income segregation in urban America has increased in recent decades (Bischoff and Reardon 2013).

Those developing indices of segregation have not come up with any satisfactory way of decomposing measures like the index of dissimilarity into what we will call "structural" and "ethnic" components, yet clearly this is vital if we are to understand the effect on racial segregation of developments like increasing income segregation. What would such a measure look like? The idea we explore in this paper is that structural segregation between two racial or ethnic groups in a metropolitan area is the amount of segregation that would remain if we could hypothetically assign households to housing units based only on their non-ethnic demographic characteristics, such as income, age, or family size. Such a simulated "race-blind" assignment of households lets us then apply any of the standard measures of ethnic segregation to the resulting demographic patterns. We draw on a new source of data and an innovative methodology to develop such simulations and analyze the indices that result. We find that our technique produces very consistent measures of structural segregation, and thus lets us infer conclusions about the "ethnic residual" – that portion of a segregation measure that is not explained by purely structural factors.

The proper interpretation and use of the measures developed here will require discussion among scholars in the field, but it seems intuitively plausible that the potential to disaggregate "structural" and "ethnic" components of housing segregation should be very useful in the realm of fair housing policy. Consider, for example, the doctrine of "disparate impact." Fair housing laws clearly prohibit housing providers from treating potential customers differently on the basis

of their race, or other protected characteristics (this is the "disparate treatment" doctrine). However, it has long been uncertain whether these laws also prohibit housing providers (or government agencies) from pursuing policies which have the *effect* of disproportionately excluding persons in a protected class (i.e., persons by race, family status, etc.) without a compelling business justification. This question was partially settled in June 2015, when the Supreme Court ruled that these "disparate impact" cases are indeed cognizable under federal fair housing law,[3] but the Court noted that it is often difficult to determine whether practices that have a nominal disparate racial impact are in fact harmful. One way of thinking about this question, in racial contexts, is whether a policy tends to increase or decrease the level of *structural* segregation among racial groups. In other words, it is helpful to be able to quantify what impact the cost of housing, the size of units, and other factors, interacting with the location of housing, has upon the *capacity* of different racial groups to achieve housing integration. Such questions, as we will see in the discussion section below, are more amenable to substantive analysis when we can disaggregate the structural component of conventional segregation measures from the ethnic component of these measures.

**Background**

To effectively measure housing segregation, one needs detailed locational data on how households are distributed across an urban area. In the United States, the dominant source of

[3] The case was intriguing in part because the plaintiff, Inclusive Communities, argued that certain subsidized housing projects in Dallas were *over-concentrated* in minority areas (Texas Department of Housing and Community Affairs v. Inclusive Communities Project, Inc., 576 U.S. 2015).

such data is the U.S. Census. Since 1940, the Census has published racial counts and other socioeconomic tabulations for small geographic units called census tracts.[4] These tracts – which vary widely in size and population, but average about four thousand residents – are the standard unit for most neighborhood-level analyses by social scientists. Most measures of segregation are based on census tracts as well.[5]

One common strategy for getting at the "ethnic" vs. "structural" question is to examine the level of racial housing segregation across specific income ranges (e.g., among black and white households with annual incomes between $50,000 and $75,000). Such studies tend to find that black/white segregation levels within a given metropolitan area are nearly as high when one focuses on a specific income range as when one compares the entire black and white populations (Farley 1995; Fischer 2003; Massey and Denton 1993; Massey and Fischer 1999). The same pattern generally holds true when one measures segregation between blacks and whites in a single educational band (e.g., households where the head has a bachelor's degree but no graduate degree) (Iceland et al. 2005). The implication drawn from these patterns is that "ethnic" factors dominate "structural" factors in explaining African-American segregation. In contrast, similar analyses have found that among Hispanics, segregation from whites depends significantly upon

[4] The census started reporting tract data for New York City in 1910 and gradually expanded the practice to other cities. Most major cities were tracted by 1940. All metropolitan areas were fully tracted by 1970 and the entire United States was tracted by 1990.

[5] Racial counts (but very little other demographic data) have also been available for several decades at the block level, which averages around 250 residents. Segregation indices tend to rise as the level of resolution increases, so block-level measures of racial segregation tend to be about .05 points higher than tract-level measures.

socioeconomic status (more affluent Hispanics are substantially more integrated) and upon whether one is a recent immigrant, a second-generation native, or a multi-generational native – all of which suggests that "structural" factors play a larger role for Hispanics (Iceland and Scopilliti 2008).

Another line of research finds that levels of racial housing segregation are strongly influenced by the degree of exclusionary zoning in a metropolitan area. For example, Rothwell (2011) uses data on land regulation across fifty metropolitan areas to derive a measure of density zoning.[6] He finds that the permitted level of density zoning is strongly correlated with the degree of housing segregation between whites on the one hand and "minorities" (blacks plus Hispanics) on the other, a relationship that survives a broader regression controlling for a variety of other factors. Rothwell concludes that metropolitan zoning patterns are "strongly associated and perhaps causally linked" to segregation, which would seem to imply that "structural" factors play a very large role in determining segregation levels.

Each of these approaches has significant flaws. Consider, for example, an analysis that finds that the overall level of black-white segregation (measured by the index of dissimilarity) in Metropolitan Area X is .75, and that if one only compares black and white households with incomes between $75,000 and $100,000, one gets an identical index measure (.75). The inference – that segregation in X is entirely explained by ethnic factors – is wrong for at least three reasons. First, it may be the wrong comparison. Blacks with household incomes of

[6] There are a variety of challenges Rothwell faced in doing so, and his measure only partly overcomes these, so we are skeptical whether his results provide a clear hierarchy of metro areas by the degree of density regulation. We suspect that a better method would use census data to directly calculate density variations.

$75,000-$100,000 may (and in fact, do on average) have lower wealth than white households with the same income, so may be homeowners in neighborhoods with whites having incomes in, for example, the lower $50,000-$60,000 range. Second, a feature of the index of dissimilarity is that it biases upward as the number of households shrinks, so the small number of black households in the $75,000-100,000 income range will create a measurement error in this comparison (Fossett in press; Winship 1977). For the same reason, the SES comparison approach breaks down when one tries to control for multiple SES characteristics: the comparison groups become far too small for meaningful analysis. Last but not least, no student of segregation would contend that structural segregation is non-existent; the SES comparison approach may tell us that ethnic segregation is important, but it has no capacity to tell us much about the scale of structural segregation, or how the structural and ethnic components compare across urban areas or across decades.

The Rothwell (2011) approach is even more flawed. His analysis suggests that variations in structural segregation across metropolitan areas can account for much of the variation in desegregation rates across these areas – which of course implies that structural factors are dominant in explaining racial segregation. However, the correlations Rothwell produces with levels of segregation cannot be taken at face value. First, as Rothwell's supplementary analyses indicate, changes in structural factors over time have very little power in predicting changes in (rather than the levels of) residential segregation. Second, the degree to which segregation has fallen across urban areas since 1970 correlates with many different race-related factors, such as black inter-metropolitan migration, the presence of key white European ethnic groups in a neighborhood, and white demand for housing in integrated neighborhoods. Several of these can be shown to have a direct effect on changes in segregation over time (Sander 1998). Third, since

zoning is primarily a local phenomenon, measuring the average level of exclusionary zoning – as Rothwell does – obscures intra-metropolitan zoning variation. Most fundamentally, the Rothwell approach does not let us disaggregate "ethnic" and "structural" factors, and, like the SES comparison method, does not allow us to quantitatively decompose the impact of structural factors and ethnic factors upon segregation.

**Analytical Strategy, Methods, and Data**

Our goal in this section is to outline an approach for computing a pure measure of structural segregation across racial or ethnic lines.

Suppose that we had complete information on every household, including their demographic characteristics, their housing characteristics, and their precise location within a metropolitan region. We could then determine how each household has translated their demographic characteristics (e.g., their income, age, household size) into some particular package of housing characteristics. Each occupied housing unit could thus become a bin, representing the housing bundle that a household with some particular mix of demographic characteristics has chosen.

This is the basic insight behind our methodology. Each occupied housing unit is tagged with the key demographic characteristics of its current household. Then, through a simulation, all housing units are emptied and their households are mixed in a large beaker. A household is drawn from the beaker and assigned to a housing unit whose key demographic characteristics match. That becomes the household's simulated location. The same process is repeated for every other household in the metropolitan area. What we end up with is a distribution of households that is determined only by their non-racial demographic characteristics. If we compute the index

of dissimilarity among, say, black and white households in our simulated urban area, we have a good estimate of the upper-bound of black/white integration given the structural distribution of housing across the region.

Performing this type of calculation requires detailed household-level data, allowing the researcher to identify all the reported demographic and housing characteristics of a given household. The Census has long released Public Use Microdata Samples (PUMS), but until 1990 the most detailed geography available in PUMS were large cities, urban counties, or states. Since 1990, the most detailed geographic unit has been the Public Use Microdata Area (PUMA), each of which has a population of at least 100,000. Using units this large to measure metropolitan segregation will almost always produce substantial underestimates. Consequently, the introduction of the Federal Statistical Research Data Center (RDC) program – which allows licensed researchers to examine all collected census data from recent decades without geographic aggregation – is crucial to fully implement our methodology.

We decided to do our analysis at the census tract level, mainly because this is the geography most widely used for calculating segregation indices. A more difficult question concerned what demographic markers to use in creating bins for the households and the housing units. Our goal was to identify a few variables, and a few ranges within each of those variables, that are important in determining the physical type of housing unit that adults will seek. We needed to be parsimonious in choosing these variables and ranges, because every combination of each variable-range with other variable-ranges would become a new sorting bin. With this in mind, we arrived at the following:

*Household Income, Normalized to the Metropolitan Median*. Most housing research, as well as common sense, suggest that income is the single more important determinant of the

amount of housing households purchase and consume. We used six income ranges, cast in terms of the metropolitan median household income to effectively normalize our ranges to the time and urban location of each analysis. Our ranges were: (a) less than 50% of the metropolitan median income, (b) 50-99%, (c) 100-149%, (d) 150-199%, (e) 200-299%, and (f) 300% or more. We skewed our bins upward (i.e., rather than having three bins below and three bins above median income) because income distributions are themselves skewed upwards, and we did not want to understate the problem of exclusionary zoning, which is thought to particularly isolate the affluent.

*Household Size*. The second-most important determinant of housing consumption appears to be household size, with larger households of course living in larger units, when one adjusts for other demographic characteristics. We divided households into three size ranges: 1 person; 2 to 4 persons; and 5 or more persons.

*Age of householder*. Age also importantly determines both the type of housing and the type of neighborhood that householders seek, with (to simplify) younger households generally seeking dense locations and older households generally seeking low-density locations. We used four age categories for the householder: (a) 29 years of age and under, (b) 30 to 39, (c) 40 to 54, and (d) 55 and older.[7]

*Presence of children*. Whether a household has children can be strongly predicted from its household size and the age of its householder, but when the prediction is incorrect, the

[7] Note that fair housing laws allow for the development of "senior" communities that can exclude children and householders under the age of 55.

housing patterns of households with and without children vary significantly. We thus divided households into those that did or did not include one or more minor children.[8]

*Tenure*. All of the variables noted above are strictly demographic and economic identifiers of households. The general strategy of our method is to let these characteristics define the housing unit they chose, rather than using housing characteristics themselves, but in the end we decided to include tenure – i.e., whether the household owned or rented its unit. The most important reason for doing so is that blacks and Hispanics have substantially different rates of home ownership compared to whites, even when one controls for household income. One key reason for this is that black and Hispanic households have, on average, substantially less wealth than white households with similar levels of income. This makes it harder for a given black or Hispanic household to afford to buy a house, and also often creates a feedback mechanism in which lack of homeownership reduces wealth accumulation. Historic patterns of segregation also tend to concentrate black and Hispanic households in the denser parts of metropolitan areas, where homeownership is less common. To the extent that discriminatory barriers reduce minority access to housing markets, they will also militate against minority ownership.

By including homeownership as a bin variable, therefore, we attempt to adjust for all of these otherwise hard-to-quantify factors. Including homeownership inflates somewhat our measure of structural segregation. To see an example of this more concretely, imagine that all income and other socioeconomic differences across blacks as a group and whites as a group could be suddenly eliminated. The residual effect of past segregation would still mean that blacks were far more concentrated than whites near the center of metropolitan areas, and therefore more likely, other things being equal, to be renters rather than owners, because they

[8] A “minor” child is one under the age of 18.

would live in more densely-populated portions of metro areas. Our inclusion of owner/renter bins thus treats as "structural" something that at least partly reflects racial and ethnic differences in location. We consequently have also produced sensitivity estimates without this variable. We choose, however, to present our main results with the tenure variable included in the simulation since it serves as a proxy for wealth, a variable not available in the U.S. Census.

These six variables, and their individual possible values, generate a total of 288 possible combinations. Our methodology, then, can be simply restated:

--Identify a metropolitan area of interest;

--Tag every occupied housing unit in the long-form census sample with its bin number and its census tract;

--Tag every household in the long-form census sample with its bin number, preserving identification of the household's race and/or ethnicity;

--Mix the households thoroughly and randomly, sequentially assign each household to a housing unit in the same bin.

The new set of reassigned households allows us to measure structural segregation along any dimension we like; we can examine the simulated distribution of black householders and compare them with white householders or we can compare Hispanic and non-Hispanic white households.

*Data*

As noted earlier, the data used in this exercise is in-house census data available through the Census Bureau's Federal Statistical Research Data Center (RDC) program. We made use of the long-form Decennial Census files for 1960, 1970, 1980, 1990, and 2000, and the American

Community Survey (ACS) files for 2006-10. Before the onset of the ACS, the Decennial Census attempted to survey every household in the nation with either a short form (sent to every household) or a long form. The short-form questionnaire was limited to an enumeration of the number of persons in each household, and the race, age, and relationship to the household head of each person. The long-form questionnaire, sent to a sample of roughly one-sixth of all households, asked dozens of additional questions about the education, occupation, earnings, citizenship, fertility, and other characteristics of each household member, as well as many questions about the household's housing. We use data from the long-form because we need data on household income in addition to demographic data on household size, presence of children, and the age of the householder. We disregard the population in group quarters and any population outside of metropolitan areas.

In 2010, the Decennial Census dispensed with the long form, focusing instead on a complete short-form count of the total population. Replacing the long form was the ACS, which substantively covers nearly the same topics that were formerly on the long form. The great advantage of the ACS is that, since it is conducted annually, we have available at any given time (with a two-year lag) fairly current data on American households, and it is possible to understand changes that occur between census years in a way we formerly could not. The disadvantage of the ACS is that each year's sample is far smaller (about 1% of all households) than the old long-form data (15-20% of all households) from the decennial census. This is a particular problem when one is trying to estimate demographic characteristics of a small area, such as a census tract.[9] For these reasons, some caution should be used in comparing tract-level analyses – such as our simulations – of the 2006-10 ACS with analyses of earlier, decennial long-form data.

[9] If a typical tract has 4,000 residents, then a 1% sample produces only 10-15 households.

*Index of Dissimilarity*

After randomly assigning each household in a metropolitan area to a different housing unit matching on household size, householder age, presence of children, tenure, and income, we recalculate the index of dissimilarity at the census tract level using the new geographic distribution of the population of households. The index of dissimilarity is defined by the following formula:

$$D = 1/2 \sum \left| \frac{N_{1i}}{N_1} - \frac{N_{2i}}{N_2} \right|,$$

where $N_{1i}$ is the population of group 1 in the *i*th tract, $N_{2i}$ is the population of group 2 in the *i*th tract, $N_i$ is the total population of group 1 in the metropolitan area, and $N_2$ is the total population of group 2 in the metropolitan area. The index value can be interpreted as the proportion of group 1 (or group 2), which would have to move to a different census tract so that each census tract would have the same composition as the metropolitan area as a whole (White 1983).

Our estimate of the "structural" segregation within a metropolitan area is the index of dissimilarity of the randomly distributed population of households. Our estimate of the "ethnic" segregation within a metropolitan area is the difference between the actual household-level index of dissimilarity and the estimated household-level index of dissimilarity of the randomly distributed population of households.

*Metropolitan area definitions*

Since we make comparisons between and within metropolitan areas over time, we have standardized the size of all metropolitan areas using the 1980 official Standard Metropolitan Statistical Area (SMSA) boundaries. This approach allows us to compare how the "structural"

and “ethnic” components of segregation vary over time using consistent geographic units of analysis. It does, however, have the disadvantage of excluding newer suburban areas from our analyses. To the extent that newer suburban enclaves house mostly white affluent residents, we may underestimate the “structural” component of segregation.

*Standard errors*

In addition to computing point estimates of our simulated structural segregation indexes, we also run each simulation 100 times to produce standard errors of our randomization procedure. All standard errors are very small.

**Results**

African-Americans. Both the simplest and the most striking result from our simulation analysis is that structural segregation between blacks and non-blacks, as measured by the index of dissimilarity, is relatively low: generally in the neighborhood of .2 on a scale where 0 represents complete integration and 1 represents complete segregation. In other words, if blacks and whites were randomly assigned to existing housing units in their metropolitan areas, and no change was made in the distribution or cost of housing, in the proportion of blacks and whites owning their own homes, or in the incomes of households, then the level of segregation between blacks and whites would be in the neighborhood of .2. In the parlance of segregation scholarship, this is an extraordinarily low measure. Real-world segregation of blacks in urban areas, over the period we are examining, has generally ranged from .5 to .9, with most blacks over most of this period living in environments at the top of this range. Some scholars have suggested that

dissimilarity values of 0 to .3 reflect low segregation, vales of .3 to .6 reflect moderate segregation, and values of .6 or above reflect high segregation. In reality, values below .3 are rarely observed in America, or even in Western Europe, between any two well-defined ethnic groups, even when there is little public awareness of any difference in residential patterns. For example, Chinese-Americans are now considered well-integrated in most metropolitan areas where they reside in significant numbers, but the index of dissimilarity between whites and Chinese-Americans is consistently over .3 and often above .4. Jewish-Americans (at least, if we use the best available census proxy for Jews, and count Russian-Americans) also have, relative to whites, an index of dissimilarity in the .3 to .4 range. Even very assimilated white groups, such as Polish-Americans, generally have an index of dissimilarity with other whites in the .2 to .3 range. Thus, the finding that there is no structural barrier to achieving even greater levels of integration for blacks is quite significant.

[Table 1 about here]

Table 1 shows how the black/white index of structural dissimilarity evolved over the period from 1960 to 2010 in the twenty-five metropolitan areas with the largest black populations and compares these with the actual dissimilarity levels for these periods. The structural indices declined sharply in the 1960s, more gradually and irregularly from 1970 to 2000, and then rose significantly between 2000 and 2010. The decline in the 1960s is consistent with other evidence from that period: housing stock expanded dramatically in the 1960s, and blacks converged with whites during that decade on all the measures we used in our binning process: income, age distribution, and family size. The rise in structural segregation in the 2000-2010 decade also comports with other data. Notably, Bischoff and Reardon (2013) have shown that economic residential segregation increased markedly during this decade.

Table 1 also shows, in the right column, the difference between the "real" and the "structural" black/white dissimilarity. We call this the "ethnic" component of dissimilarity. In other words, "structural" dissimilarity reflects that portion of segregation that directly reflects the facts that housing varies widely across a metropolitan area and any two ethnic groups we might compare will have different demographic makeups that lead them to occupy different types of housing and thus create "structural" differences in residential location. The difference between these structural levels of dissimilarity and the actual, observed levels reflect mechanisms that are directly or indirectly tied to race or, more broadly, ethnicity. As Table 1 makes clear, the general declines in segregation that have occurred since 1970 are entirely accounted for by declines in the "ethnic" component of segregation, since the "structural" component, despite its initial decline, is higher in 2010 than in 1970.

[Table 2 about here]

The minor role played by structural change is even clearer when we consider the patterns in Table 2. Table 2 shows the change in overall black/white dissimilarity, our structural measure, and the ethnic residual for individual metropolitan areas over the period 1970 to 2010. The top part of the table shows the twelve American metropolitan areas which had the largest black populations in 1980; the bottom part of the table shows the six largest metropolitan areas which had at least a ten-point decline in the index of dissimilarity during the 1970s. Among the first group, only three out of twelve metro areas experienced a decline in structural segregation over the 1970 to 2010 period, and in only one metro area (New York) was the change in structural segregation (there, a rise of 5 points) larger in magnitude than the change in the ethnic residual (a rise of 1.4 points). In the other eleven areas, the drop in the ethnic residual was larger, in both absolute and relative terms, than any change in structural segregation.

Among the metro areas that experienced substantial drops in overall segregation in the 1970s, the pattern we observe is even more striking. In all six of these areas, structural segregation actually rose quite significantly over the 1970 to 2010 period; the smallest increase in any of these areas was larger than the largest increase at any of our "big twelve" areas. What makes this particularly counterintuitive, at least at first, is that in the desegregating metro areas, black household characteristics tended to converge to white characteristics significantly more than is the case in the "big twelve," where segregation dropped more modestly and black/white differences stagnated or, in some cases, widened.

We will return to this intriguing pattern in the Discussion section, below. For now, we will note that, because of the consistent pattern of rising structural segregation in these metro areas, and because all of these areas had substantial drops in gross black/white dissimilarity, the declines in the ethnic residual in these six areas were particularly large, from twenty to thirty-five percentage points. In other words, the ethnic residual in all six of these areas fell by more than half from 1970 to 2010 – extraordinary declines in major metropolitan areas.

Everything in Table 2 suggests, at least at first inspection, that to the extent that black/white segregation has declined during the Civil Rights Era, the declines have been driven by changes related to the "ethnic" component of this segregation rather than "structural" factors. The implication – though this is an implication that requires much closer examination – is that gains in black income, or the gradual convergence of black/white household sizes, over the past half-century, have been less important in shaping the decline of segregation than have changes related to race – such as declines in racial discrimination and an increased willingness to live in integrated neighborhoods.

Hispanics. The other major minority group in the United States are Hispanics, and they present a particularly interesting contrast with African-Americans. Hispanic segregation has generally been seen as following the model of ethnic Europeans rather than African-Americans: that is, new Hispanic arrivals are generally highly segregated from other groups, but this level of segregation declines steadily with length of residence in the United States, English fluency, intermarriage with non-Hispanics, and increases in income and education. On the other hand, several major metropolitan areas have witnessed steady increases in Hispanic segregation over the past generation, which at least on its face seems to conflict with the standard assimilation model.

In developing Hispanic indices comparable to the data shown for African-Americans in Tables 1 and 2, one faces a few challenges. The Census did not develop and apply the modern definition of Hispanic until the 1980 Census, so for 1960 and 1970 one must use a surrogate measure, namely Spanish mother tongue and Hispanic surname. Since Hispanics in the 1960-70 era were concentrated in the Southwest (plus Miami and New York City), comparisons of metro-level segregation levels over time are tricky; one must either look at a consistent but small set of metro areas, or use non-comparable and growing samples of cities for comparison purposes. To address this problem, and to err on the side of providing more rather than less information, we report data for a relatively large set of metro areas in Table 3. In Table 4, we present a shorter and small series of indices for a cohort of metro areas that can be compared under virtually constant standards from 1980 through 2010.

[Tables 3 and 4 about here]

Tables 3 and 4 follow the example of Table 1 by reporting real, structural, and ethnic measures of segregation between Hispanics and non-Hispanic whites. To us, the most striking

pattern in these tables is although the real level of dissimilarity between Hispanics and whites is consistently lower than that for blacks and whites, the level of structural segregation for Hispanics is generally comparable to the level for African-Americans. This necessarily means that the level of ethnic segregation for Hispanics is particularly low compared to the black level. This is consistent with the hypothesis that Hispanic segregation does follow the European ethnic model of gradual assimilation.

*Standard errors and sensitivity analyses*

As we mention above, we computed standard errors of our simulation indexes and ran several sensitivity analyses to see how different our results would be if we were to include different sets of variables in creating the bins for our simulation.

[Table 5 about here]

Table 5 shows the standard errors of our simulated dissimilarity indexes for the same sets of metropolitan areas in Tables 1 and 3. We ran our simulation procedure 100 times in order to compute the standard errors. All standard errors are very small.

In addition, we ran our simulation without using the tenure variable, with a different specification of the income variable,[10] and including the education of the household head as an additional criterion in generating our bins. All of these analyses produced very small changes in our estimates of structural segregation. Table 6 shows a summary of our results without the tenure variable. The simulated structural segregation indexes without the tenure variable are

[10] We experimented with categorizing the income variable in terms of percentiles of the household income within each metropolitan area. We used the following percentile categories: 1) 0-15; 2) 16-30; 3)31-49; 4) 50-69; 5) 70-84; 6) 85-94; 7) 95 and above.

about 1 to 3 points lower compared to the ones that include the tenure variable. These differences are not big enough to change the conclusions that we reach, but it is interesting to observe that removing the tenure variable makes a bigger difference for the Black/White structural indexes compared to the Hispanic/White structural indexes.

[Table 6 about here]

Finally, we also computed a simulated isolation index for whites, African Americans, and Hispanics. The isolation index is a well established index of segregation that measures the probability that any given member of a racial or ethnic group lives in the same neighborhood with another member of the same racial or ethnic group. Our simulated isolation indexes reveal very similar patterns to our simulated dissimilarity indexes.

**Discussion**

Urban scholars have become used to thinking about racial segregation – or other types of segregation – in terms of a single number, the overall index of dissimilarity. Our disaggregation of that number into two parts – structural segregation and the "ethnic residual" – will need to be processed by scholars and subjected to a variety of tests before a consensus emerges about the proper interpretation and use of the new numbers. However, we are confident that the numbers are meaningful, and will be quite helpful in thinking about segregation in more concrete ways.

First, as we noted in the introduction, measures of structural segregation can be helpful in evaluating cases of disparate impact. In the Supreme Court's *Inclusive Communities* decision, the state of Texas was found to concentrate certain types of moderately subsidized rental housing in minority communities. For minorities, such a policy could have both beneficial (by improving housing stock in minority communities) and harmful effects (by reducing the availability of

affordable housing in predominantly white communities). Through a structural segregation lens, we can evaluate such a policy in a more concrete and nuanced way, by considering whether the particular mix of unit cost and size in a particular development tends to increase, or decrease, overall structural segregation in the metropolitan area. Placing smaller units, priced at mid-market households, into low-and-moderate income communities, and placing less expensive units, priced at lower-income households, into middle-income communities – and making sure that within this context, particular racial groups are not singled out or excluded – could allow an agency to pursue legitimate redevelopment goals without producing any negative disparate impacts and indeed reducing structural segregation.

Another prominent concern in fair housing policy concerns exclusionary zoning – that is, zoning restrictions in urban neighborhoods or more often suburban towns that effectively prevent the construction of low-and-moderate income housing through requirements for such things as minimum lot sizes or expensive housing components. As we have noted, many fair housing advocates believe exclusionary zoning is a principal cause of racial segregation (Silver, 1997). Measures of structural segregation should make it possible to quantify the effect of particular zoning policies on overall levels of structural segregation. Such measures can help government agencies, advocacy groups, and courts in making clearer, more analytically grounded distinctions between policies that have a powerful effect upon structural segregation compared to those which have relatively marginal effects.

In a similar vein, we can see from the structural segregation measures that many metropolitan areas experienced sharp increases in structural segregation during the 2000-2010 decade. The increase may largely reflect increased income inequality. However, the existence of a structural segregation measure opens the way to examining this problem in far more concrete

ways. We can develop models, using measures of intra- and inter-racial economic disparities, as well as measures of zoning and housing changes in the metropolitan area, to determine fairly precisely just what factors are associated with increases in economic segregation. This will also let us evaluate more concretely how different proposed remedial measures might reduce structural segregation.

The ethnic residual may also generate substantial analytic dividends. An important but very difficult question concerns the mechanisms by which black/white segregation levels have fallen sharply in some metropolitan areas, while declining much more slowly in others. Often, scholars attempt to explain the path of segregation by developing models that estimate aggregate levels of segregation, and how those levels change over time. Our decomposition of the level of segregation gives analysts two distinct and more precise targets upon which to focus. Some models (e.g., the Rothwell model discussed earlier) should focus on explaining changes in structural segregation; others (by Farley, Galster, Sander and others) that are primarily concerned with the racial drivers of segregation may more fruitfully focus on the ethnic residual as a dependent variable.

As a final note, over the past few years there has been a striking upsurge in interest in the problem of housing segregation, with parallel developments in the courts, in politics, in media coverage, and among scholars. The RDC data are incredibly rich and helpful in giving our analysis of housing segregation tools it has badly needed for decades.

**Bibliography**

Bischoff, K., & Reardon, S. F. (2013). Residential segregation by income, 1970-2009. In J. R. Logan (Ed.), *The Lost Decade? Social Change in the U.S. after 2000.* New York, NY: Russell Sage Foundation. Retrieved from http://www.s4.brown.edu/us2010/Data/Report/report10162013.pdf

Farley, J. E. (1995). Race still matters: The minimal role of income and housing cost as causes of housing segregation in St. Louis, 1990. *Urban Affairs Quarterly* 31(2), 244-254.

Fischer, M. J. (2003). The relative importance of income and race in determining residential outcomes in U.S. urban areas, 1970-2000. *Urban Affairs Review* 38, 669-696.

Fossett, M. A. (In print). *New demographic techniques for measuring and investigating residential segregation*. New York, NY: Springer.

Huntington Branch NAACP v. Town of Huntington, 844 F.2d 926 (2nd Cir 1988).

Iceland J., & Scopilliti, M. (2008). Immigrant residential segregation in US metropolitan areas, 1990–2000. *Demography* 45(1), 79-94.

Iceland, J., Sharpe, C., & Steinmetz, E. (2005). Class differences in African American residential patterns in US metropolitan areas: 1990–2000. *Social Science Research* 34(1), 252-266.

Massey, D. S. & Denton, N. (1993). *American apartheid: Segregation and the making of the underclass*. Cambridge, MA: Harvard University Press.

Massey, D. S. & Fischer, M. J. (1999). Does rising income bring integration? New results for Blacks, Hispanics, and Asians in 1990. *Social Science Research* 28(3), 316–326.

Massey, D. S., & Fischer, M. J. (2000). How segregation concentrates poverty. *Ethnic and Racial Studies* 23(4), 670-691.

Muth, R. F. (1986). The causes of housing segregation. *Issues in housing discrimination*. U.S. Commission on Civil Rights, 3-13.

Quillian, L. (2002). Why is Black–White residential segregation so persistent? Evidence on three theories from migration data. *Social Science Research* 31(2), 197-229

Rothwell, J. (2011). Racial enclaves and density zoning: The institutionalized segregation of racial minorities in the United States. *American Law and Economics Review* 13, 290-358.

Sander, R. (1998). Housing segregation and housing integration: The diverging paths of urban America. *Miami Law Review* 52, 977-1010.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology* 1(2), 143-186.

Silver, C. (1997). The racial origins of zoning in American cities. In Manning T., Ritzdorf, J. & Ritzdorf, M. (Eds.) *Urban planning and the African American community: In the shadows*. Thousand Oaks, CA: Sage Publications.

Texas Department of Housing and Community Affairs v. Inclusive Communities Project, Inc., 576 U.S. (2015).

Village of Arlington Heights v. Metropolitan Housing Development Corp., 429 U.S. 252 (1977).

White, M. J. (1983). The measurement of spatial segregation. *The American Journal of Sociology* 88(5), 1008-1018.

Wilson, W. J. (1996). *When work disappears*. New York, NY: Alfred A. Knopf.

Winship, C. (1977). A revaluation of indexes of residential segregation. *Social Forces* 55(4), 1058-1066.

Table 1. Overview of Black/White Total and Structural Segregation, 1960-2010

| | Index of Dissimilarity measured through | | |
|---|---|---|---|
| Year | Actual B/W distribution | Simulated B/W distribution | Ethnic "residual" |
| 1960 | 0.801 | 0.231 | 0.570 |
| 1970 | 0.806 | 0.196 | 0.610 |
| 1980 | 0.760 | 0.182 | 0.578 |
| 1990 | 0.721 | 0.193 | 0.528 |
| 2000 | 0.701 | 0.187 | 0.514 |
| 2010 | 0.677 | 0.215 | 0.462 |

*Notes*: All MSAs are defined in constant geographic terms from 1960 through 2010, using 1980 SMSA boundaries. Indices are based on the 25 1980 SMSAs with the largest black populations. For 1960, indices are based only on the top 24 SMSAs because the Nassau-Suffolk, NY metropolitan area was not tracted.

Table 2. Black/White Patterns in Eighteen Illustrative Metro Areas

| **Metro areas: Large black population** | 1970 Index of Dissimilarity | | | 2010 Index of Dissimilarity | | | Δ D, Structural | Δ D, Ethnic |
|---|---|---|---|---|---|---|---|---|
| | Total | Structural | Ethnic | Total | Structural | Ethnic | | |
| New York | 0.742 | 0.181 | 0.561 | 0.810 | 0.231 | 0.579 | 0.050 | 0.018 |
| Chicago | 0.908 | 0.230 | 0.678 | 0.767 | 0.237 | 0.530 | 0.007 | -0.148 |
| Los Angeles | 0.887 | 0.134 | 0.753 | 0.686 | 0.223 | 0.463 | 0.089 | -0.290 |
| Detroit | 0.875 | 0.181 | 0.694 | 0.767 | 0.208 | 0.559 | 0.027 | -0.135 |
| Philadelphia | 0.781 | 0.164 | 0.617 | 0.709 | 0.212 | 0.497 | 0.048 | -0.120 |
| Washington, DC | 0.808 | 0.176 | 0.632 | 0.640 | 0.194 | 0.446 | 0.018 | -0.186 |
| Baltimore | 0.813 | 0.212 | 0.601 | 0.656 | 0.199 | 0.457 | -0.013 | -0.144 |
| Houston | 0.787 | 0.187 | 0.600 | 0.621 | 0.211 | 0.410 | 0.024 | -0.190 |
| Atlanta | 0.827 | 0.204 | 0.623 | 0.591 | 0.174 | 0.417 | -0.030 | -0.206 |
| Dallas | 0.872 | 0.191 | 0.681 | 0.574 | 0.215 | 0.359 | 0.024 | -0.322 |
| Newark | 0.773 | 0.251 | 0.522 | 0.770 | 0.297 | 0.473 | 0.046 | -0.049 |
| St. Louis | 0.843 | 0.212 | 0.631 | 0.719 | 0.194 | 0.525 | -0.018 | -0.106 |
| **Metro Areas: Large Declines in D** | | | | | | | | |
| San Antonio | 0.773 | 0.157 | 0.616 | 0.500 | 0.213 | 0.287 | 0.056 | -0.329 |
| Minneapolis | 0.787 | 0.242 | 0.545 | 0.550 | 0.294 | 0.256 | 0.052 | -0.289 |
| Seattle | 0.774 | 0.158 | 0.616 | 0.535 | 0.273 | 0.262 | 0.115 | -0.354 |
| San Jose | 0.510 | 0.199 | 0.311 | 0.449 | 0.348 | 0.101 | 0.149 | -0.210 |
| San Diego | 0.815 | 0.166 | 0.649 | 0.550 | 0.250 | 0.300 | 0.084 | -0.349 |
| Denver | 0.851 | 0.172 | 0.679 | 0.646 | 0.269 | 0.377 | 0.097 | -0.302 |

Table 3. Overview of White/Hispanic Total and Structural Segregation, 1960-2010

| Year | Index of Dissimilarity measured through | | |
|---|---|---|---|
| | Actual H/W distribution | Simulated H/W distribution | Ethnic "residual" |
| 1960 | 0.512 | 0.193 | 0.319 |
| 1970 | 0.458 | 0.169 | 0.289 |
| 1980 | 0.485 | 0.177 | 0.308 |
| 1990 | 0.475 | 0.198 | 0.277 |
| 2000 | 0.492 | 0.208 | 0.284 |
| 2010 | 0.487 | 0.218 | 0.269 |

*Notes*: All MSAs are defined in constant geographic terms from 1960 through 2010, using 1980 SMSA boundaries. Indices are based on the 25 1980 SMSAs with the largest Hispanic populations. For 1960, indices are based only on the top 21 SMSAs because the following metropolitan areas were not tracted: Anaheim-Santa Ana-Garden Grove, CA, Brownsville-Harlingen-San Benito, TX, McAllen-Pharr-Edinburg, TX, and Oxnard-Simi Valley-Ventura, CA.

Table 4. White/Hispanic Patterns in Ten Illustrative Metro Areas

| **Metro areas: Large Hispanic population** | 1970 Index of Dissimilarity | | | 2010 Index of Dissimilarity | | | Δ D, Structural | Δ D, Ethnic |
|---|---|---|---|---|---|---|---|---|
| | Total | Structural | Ethnic | Total | Structural | Ethnic | | |
| Los Angeles | 0.450 | 0.142 | 0.308 | 0.608 | 0.265 | 0.343 | 0.123 | 0.035 |
| New York | 0.619 | 0.267 | 0.352 | 0.618 | 0.286 | 0.332 | 0.019 | -0.020 |
| Chicago | 0.556 | 0.233 | 0.323 | 0.550 | 0.244 | 0.306 | 0.011 | -0.017 |
| Miami | 0.506 | 0.159 | 0.347 | 0.523 | 0.197 | 0.326 | 0.038 | -0.021 |
| San Antonio | 0.569 | 0.184 | 0.385 | 0.459 | 0.160 | 0.299 | -0.024 | -0.086 |
| Houston | 0.422 | 0.142 | 0.280 | 0.505 | 0.216 | 0.289 | 0.074 | 0.009 |
| San Francisco | 0.328 | 0.118 | 0.210 | 0.452 | 0.221 | 0.231 | 0.103 | 0.021 |
| El Paso | 0.479 | 0.173 | 0.306 | 0.430 | 0.200 | 0.230 | 0.027 | -0.076 |
| Riverside | 0.378 | 0.123 | 0.255 | 0.408 | 0.193 | 0.215 | 0.070 | -0.040 |
| Anaheim | 0.304 | 0.116 | 0.188 | 0.502 | 0.250 | 0.252 | 0.134 | 0.064 |

Table 5. Standard Errors of the Simulated Segregation Index, 1960-2010

| | Standard error for | |
|---|---|---|
| | Simulated B/W distribution | Simulated H/W distribution |
| 1960 | 0.0003 | 0.0006 |
| 1970 | 0.0003 | 0.0005 |
| 1980 | 0.0003 | 0.0005 |
| 1990 | 0.0003 | 0.0005 |
| 2000 | 0.0003 | 0.0004 |
| 2010 | 0.0005 | 0.0005 |

*Notes*: All MSAs are defined in constant geographic terms from 1960 through 2010, using 1980 SMSA boundaries.

Table 6. Simulated Structural Segregation Index without Tenure, 1960-2010

| | Index of Dissimilarity measured through | |
|---|---|---|
| | Simulated B/W distribution | Simulated H/W distribution |
| 1960 | 0.188 | 0.178 |
| 1970 | 0.164 | 0.159 |
| 1980 | 0.157 | 0.165 |
| 1990 | 0.169 | 0.188 |
| 2000 | 0.159 | 0.195 |
| 2010 | 0.191 | 0.207 |

*Notes*: All MSAs are defined in constant geographic terms from 1960 through 2010, using 1980 SMSA boundaries.