Characteristics of the Joint Distribution of Race and Income Among Neighborhoods

Sean F. Reardon
Joseph Townsend
Lindsay Fox
Stanford University

Paper prepared for
The Russell Sage Foundation Journal Conference:
The Spatial Foundations of Inequality
February 12, 2015
New York City

DRAFT: January 19, 2015

Direct correspondence to Sean F. Reardon, Graduate School of Education, Stanford University, 520 Galvez Mall, #526, Stanford, California 94305. Phone: 650-736-8517. E-mail: sean.reardon@stanford.edu.
Introduction

Although racial and socioeconomic segregation are persistent features of the residential landscape, both have changed in the last four decades in the United States. Racial segregation has declined moderately, particularly segregation between white and black households, but remains very high in many places (Glaeser & Vigdor, 2012; Logan & Stults, 2011; Logan, Stults, & Farley, 2004). Segregation by income has risen sharply since 1970; most of that increase occurred in the 1980s and 2000s (Bischoff & Reardon, 2014; Jargowsky, 1996, 2003; Reardon & Bischoff, 2011; Watson, 2009). Less clear, however, are the trends and patterns of the joint distribution of neighborhoods’ racial and socioeconomic characteristics. That is, we do not have a clear description of how much neighborhoods differ (in terms of racial and economic composition) among households with the same income but that differ by race, or among those of the same race but differing by income. Without such a description, it is unclear whether and how changes in racial and economic segregation have altered disparities in neighborhood conditions.

In this paper, we demonstrate a general approach to describing the joint distribution of race and income among neighborhoods. We are not the first to suggest methods of describing features of this joint distribution, but our approach is one that is much more general and versatile than others. We show that estimating a set of multidimensional exposure functions is sufficient to generate a wide range of useful statistics regarding the joint distribution of racial and economic composition of neighborhoods, including many of the measures proposed and used in a more ad hoc fashion in much of the literature.

Our interest in developing these methods derives from the theoretical and empirical literature describing the ways in which neighborhoods affect their residents’ educational, socioeconomic, and health outcomes. In much of the neighborhood effects research, neighborhood poverty (or socioeconomic conditions more generally) is hypothesized to be a key distal driver of neighborhood effects (Acevedo-Garcia & Lochner, 2003; Brooks-Gunn, Duncan, & Aber, 1997; Leventhal & Brooks-Gunn, 2000; Sampson, 2008; Sampson, Raudenbush, & Earls, 1997; Wodtke, Harding, & Elwert, 2011). A
neighborhood’s income distribution is hypothesized to directly or indirectly affect housing conditions, school and child-care quality, access to healthy food, green spaces, safe playgrounds, social networks, the prevalence of adult role models, and a range of other institutional and collective resources that are beneficial for child development.

Although some research has suggested that neighborhoods have no significant effect on many aspects of children’s development, educational success, and social, behavioral, or economic outcomes (Kling, Liebman, & Katz, 2007; Ludwig et al., 2013; Sanbonmatsu, Kling, Duncan, & Brooks-Gunn, 2006), recent rigorous research suggests that neighborhood socioeconomic conditions can have substantial effects on such outcomes, particularly as a result of sustained exposure during childhood (Burdick-Will et al., 2011; Chetty, Hedren, & Katz, 2015; Harding, 2003; Wodtke et al., 2011). Indeed, Chetty et al.’s (2015) analysis of the Moving to Opportunity experiment shows that children in families who used a (randomly assigned) housing voucher to move to a low-poverty neighborhood have substantially higher college attendance rates and 31% higher earnings by their mid-twenties than those in a control group not assigned a voucher, a finding that suggests that neighborhood poverty (or its correlates) is harmful to young children’s development.

Given these findings, and the theoretical importance of neighborhood composition, we develop and demonstrate in this paper a highly general approach to measuring the joint patterns of racial and economic neighborhood composition. In effect, this approach relies on estimating the average race-specific income distribution in the neighborhoods of individuals of any specific income and race/ethnicity. The functions describing these distributions can be used for a wide range of types of descriptive analyses, and provide a detailed account of the joint distribution of race/ethnicity and income across neighborhoods.
Measuring Segregation

Hundreds of articles have been devoted to developing and describing ways of measuring racial and economic segregation; hundreds more are devoted to describing their trends (e.g., James & Taeuber, 1985; Massey & Denton, 1993; Massey & Denton, 1988; Reardon & Bischoff, 2011; Reardon & Firebaugh, 2002; Reardon & Owens, 2014). The welter of methodological approaches to measuring segregation is partly due to academics’ penchant for methodological hair-splitting. But there are also important theoretical and conceptual distinctions about the features of segregation that are important to measure and understand, and these distinctions lead to different measurement approaches.

Despite the abundance of ways of measuring segregation, most approaches are limited to measuring segregation along a single population dimension at a time. We know, for example, how to measure segregation among two or more racial groups (James & Taeuber, 1985; Massey & Denton, 1988; Reardon & Firebaugh, 2002); among ordered occupational or educational groups (Reardon, 2009); and by income or any other single continuous dimension (Jargowsky, 1996; Reardon, 2011; Reardon & Bischoff, 2011). Methods of measuring multidimensional patterns of segregation, such as the joint distribution of race and income among neighborhoods, however, are less well-developed.

Two approaches have been used to describe features of the joint distribution of racial and economic segregation patterns. One approach measures racial segregation among households of similar income (Adelman, 2004; Darden & Kamel, 2000; Denton & Massey, 1988; J. E. Farley, 1995), or income segregation among households of the same race (Bischoff & Reardon, 2014; R. Farley, 1991; Jargowsky, 1996; Reardon & Bischoff, 2011). Typically, this approach relies on evenness measures of segregation, such as the dissimilarity index or similar measures (James & Taeuber, 1985). Some of these studies allow the comparison of overall racial segregation levels with within-income category segregation levels as a means of testing hypotheses about the role of income in shaping racial segregation levels, but they do not provide a clear description of the joint distribution of race and income across neighborhoods.
A second approach relies on so-called “exposure” measures of segregation to describe the average extent to which households of a given race-by-income category are exposed to those of another category. Most commonly, these studies compute different racial groups’ exposure to poverty: the average proportion of poor residents in the neighborhoods of members of different racial groups (Logan, 2002, 2011; Timberlake, 2002, 2007; Timberlake & Iceland, 2007). These measures provide a much more interpretable description of differences in average neighborhood socioeconomic conditions than the evenness measures. Logan (2011), for example, categorizes households by race and three income categories (poor, middle, and affluent) and measures their exposure to their own racial group and to whites. This allows him to make descriptive statements like “Affluent blacks are currently less [exposed to black neighbors] than poor blacks (36.3 vs. 42.9), and also somewhat more exposed to whites (42.9 vs. 39.8).”

The drawback of these measures is that, unless fine-grained income information is available, they are limited to comparisons based on the income categories reported in the data, which may be relatively crude. Moreover, these income categories may change over time (e.g., the Census or American Community Survey (ACS) often change the number and location of the income categories they use in published tables). Even if they don’t change over time, their location in the income distribution will vary across time and place, because of differences in income distributions. Finally, even within a given place and time, the categories are not necessarily exactly comparable across groups. Suppose we define poor as having an income below $20,000. By this definition, the average “poor” black household will generally have a lower income than the average “poor” white household, simply because the black income distribution is lower than the white distribution. So a comparison of the neighborhoods of poor whites and blacks may be misleading when we base the comparison on broad income categories rather than exact income.

We adopt an approach similar to the second one here, but develop methods of estimating the
joint distribution of racial and economic neighborhood composition in ways that are not sensitive to the
definitions of income categories provided in the Census or ACS. Our approach allows one to describe 1) the average racial and income distributions in neighborhoods of households of different incomes, 2) the average racial and income distributions in the neighborhoods of households of different income levels and race/ethnicity, and 3) the average race-specific income distributions in the neighborhoods of households of different income levels and race. These measures are similar to the more standard exposure measures used by others (e.g., the exposure of poor blacks to middle-class neighbors), but differ in that they are fully continuous, rather than categorical, measures of exposure. In effect, they describe the average joint distribution of race and income in the neighborhoods of individuals of any specific income and race/ethnicity. As a result, they can be used for a wide range of types of descriptive analyses, and provide a detailed account of the joint distribution of race/ethnicity and income across neighborhoods.

**Estimating Average Neighborhood Income Distributions, by Race and Income**

**Notation**

In this paper, we use \( g \) and \( h \) to denote racial groups (or other categorical groups); we use \( p \) and \( q \) to denote income levels, expressed as percentiles (scaled from 0 to 1, for convenience) of the population income distribution; and we use \( i \) to index neighborhoods. The function \( \rho_x(p) \) describes the income density function in some population \( x \), where \( x \) may refer to a specific group and/or neighborhood; correspondingly, the function \( R_x(p) = \int_0^p \rho_x(r)dr \) denotes the cumulative income distribution function in population \( x \). Finally we use \( T_x \) to denote the count of households in population \( x \), and \( \pi_{gi} \) to indicate the proportion of households in neighborhood \( i \) that are in group \( g \).

**Estimands**

Our goal, in general, is to estimate the function \( f^{gh}(p, q) \) that describes the average cumulative
income distribution function of group $h$ in the neighborhoods of members of group $g$ with income $p$.

That is

$$ f_{g}^{h}(p, q) = \sum_{i} \left[ \frac{T_{i} \cdot \pi_{gi} \cdot \rho_{gi}(p) \cdot T_{i} \cdot \pi_{hi} \cdot R_{hi}(q)}{T_{i}} \right]. $$

(1)

Note that $f_{g}^{h}(p, q)$, defined this way, is interpreted as the weighted average proportion of the households in a neighborhood that are members of group $h$ with incomes less than or equal to $q$, where the weights are the number of households of group $g$ with income $p$ in each neighborhood. In the segregation literature, such measures are called “exposure” measures, because they describe the average extent to which members of one group (in this case members of group $g$ with income $p$) are exposed to members of another group (members of group $h$ with incomes less than or equal to $q$) in their local context (neighborhood in this case) (Lieberson, 1981; Massey & Denton, 1988).

If we know $f_{g}^{h}(p, q)$ for all groups $g$ and $h$, we can derive a number of additional useful quantities:

Functions describing exposure to overall (not race-specific) neighborhood income distributions, conditional on race and income. The average cumulative income distribution in the neighborhoods of members of group $g$ with incomes $p$ is simply the sum of the corresponding group-specific functions:

$$ f_{g}^{t}(p, q) = \sum_{h} f_{g}^{h}(p, q). $$

(2)

Functions describing exposure to race-specific neighborhood income distributions, conditional on income. The typical household with income $p$ lives in a neighborhood where members of group $h$ have an income distribution given by

$$ f_{t}^{h}(p, q) = \frac{1}{p} \sum_{g} \left[ \pi_{g} \cdot \rho_{g}(p) \cdot f_{g}^{h}(p, q) \right]. $$
The function describing exposure to overall neighborhood income distributions, conditional on income. Combining (2) and (3) above, we can derive the function $f^t_t(p, q)$ which describes the average cumulative income distribution function in the neighborhoods of households with income $p$:

$$f^t_t(p, q) = \sum_h f^h_t(p, q) = \frac{1}{p} \sum_g \pi_g \rho_g(p) f^h_g(p, q) = \frac{1}{p} \sum_g \pi_g \rho_g(p) \sum_h f^h_g(p, q).$$

(4)

Average neighborhood racial composition, conditional on race and household income. The average racial composition of the neighborhoods of members of group $g$ with income $p$ (the exposure of members of group $g$ and income $p$ to members of group $h$) is simply $f^h_g(p, 1)$. The average racial composition in the neighborhoods of households with income $p$ is likewise given by the functions $f^h_t(p, 1)$. Note that $f^t_t(p, 1) = f^t_t(p, 1) = 1$ by definition (because all households in a neighborhood have incomes less than or equal to 1 by definition).

Average neighborhood race-specific income density functions, conditional on race and income. Because $f^h_g(p, q)$ is a cumulative distribution function, we can obtain the corresponding density function, denoted $\rho^h_g(p, q)$, by taking the derivative of $f^h_g$ with respect to $q$:

$$\rho^h_g(p, q) = \frac{d}{dq} f^h_g(p, q).$$

(5)

The formula holds when $g$ and/or $h$ is replaced by $t$ as well.

Percentiles of average neighborhood race-specific income distributions, conditional on race and income. First define

$$f^{*h}_g(p, q) = \frac{f^h_g(p, q)}{f^h_g(p, 1)}.$$

(6)
Now $f^{*h}_{g}(p, q)$ describes, for members of group $g$ with incomes $p$, the weighted average proportion of the households in a neighborhood that are members of group $h$ with incomes less than or equal to $q$ relative to the weighted average proportion of the households in a neighborhood that are members of group $h$, where the weights are defined as above. If we wanted to know the median income among the group $h$ neighbors of a member of group $g$ with income $p$, we would find $q$ such that $f^{*h}_{g}(p, q) = 0.50$.

More generally, the $100 \cdot c^{th}$ percentile of the income distribution of members of group $h$ in the neighborhoods of members of group $g$ with income $p$ is $f^{*h-1}_{g}(p, c)$, where $f^{*h-1}_{g}(p, c)$ returns the value $q$ such that $f^{*h}_{g}(p, q) = c$.

Note that $f^{*t}_{g}(p, 1) = f^{t}_{g}(p, 1) = 1$ by definition, but $f^{*h}_{c}(p, 1) \neq 1$ in general, so

\[
\begin{align*}
    f^{*t}_{g}(p, q) &= f^{t}_{g}(p, q) \\
    f^{*t}_{c}(p, q) &= f^{t}_{c}(p, q) \\
    f^{*h}_{c}(p, q) &= \frac{f^{h}_{c}(p, q)}{f^{h}_{c}(p, 1)}
\end{align*}
\]

(7)

Standardized measures of between-group differences neighborhood income distributions, conditional on household income. We might want to measure the difference between the average neighborhood income density functions for two groups $g_1$ and $g_2$, conditional on $p$; that is, for any given value of $p$, we want to measure the difference between the distributions $\rho^{t}_{g_1}(p, q)$ and $\rho^{t}_{g_2}(p, q)$. We could do this by measuring, for example, the difference in their medians (i.e., by comparing $f^{*t-1}_{g_1}(p, .50)$ and $f^{*t-1}_{g_2}(p, .50)$), but this would not provide a summary measure of the overall difference in the distributions. A useful summary measure of the degree of overlap of two distributions is the probability that a randomly chosen value from one distribution is larger than a randomly chosen value from the other. In our case here, this is the probability that a member of the neighborhood of the typical group $g_1$ household with income $p$ has an income higher than that of a randomly chosen member of the
neighborhood of the typical group \( g_2 \) household with income \( p \). This probability is equal to

\[
Pr_{g_1 > g_2}(p) = \int_0^1 f_{g_1}^{-1}(p, f_{g_2}^{-1}(p, c)) dc.
\]  

(8)

This probability can be converted to the \( V \) statistic, a non-parametric measure of the difference between two distributions:

\[
V_{g_1 g_2}(p) = \sqrt{2} \Phi^{-1}(Pr_{g_1 > g_2}(p)),
\]  

(9)

where \( \Phi^{-1}(\cdot) \) is the probit function. Here \( V_{g_1 g_2}(p) \) is a function of \( p \) that describes the extent of overlap between the typical neighborhood income distributions. \( V \) can be interpreted as the standardized difference between the means of two normal distributions with the same degree of overlap as the distributions of interest, so it is interpretable as a “pseudo effect size” (Ho & Haertel, 2006; Ho & Reardon, 2012; Holland, 2002).

**Standard exposure measures.** We can obtain additional exposure measures, such as the exposure of members of group \( g \) with incomes between \( p_{\min} \) and \( p_{\max} \) to members of group \( h \) in with incomes between some \( q_{\min} \) and \( q_{\max} \), by computing

\[
\int_{p_{\min}}^{p_{\max}} \int_{q_{\min}}^{q_{\max}} \rho_g(r) f_g^h(r, q) \rho_g^{-1}(r) dr dt.
\]  

(10)

A useful special case of this is the exposure of those in group \( g \) with income less than or equal to \( p \) to those in group \( h \) with income less than or equal to \( q \). Denoted \( F_{g h}^h(p, q) \), this is

\[
F_{g h}^h(p, q) = \frac{\int_p^q \rho_g(r) f_g^h(r, q) dr}{\int_0^p \rho_g(r) dr}
\]  

(11)
For example, the exposure of group $g$ to poor neighbors would be $F^t_g(1, q_{poverty})$, where $q_{poverty}$ is the income value that corresponds to the poverty line. Thus, measures of “exposure to poverty” used in much of the segregation literature (Logan, 2011; Timberlake, 2002) are special cases of the measurement approach we describe here. Note that in the special case where $p = q = 1$, $F^h_g(1,1)$ is a standard exposure measure of racial segregation, the exposure of group $g$ to group $h$ (usually denoted $gP^*_h$). In our notation, this standard exposure measure can be written

$$gP^*_h = F^h_g(1,1) = \int_0^1 \rho_g(r)f^h_g(r,1)dr.$$  

Thus, estimating $f^h_g(p, q)$ is sufficient to obtain a number of useful functions describing the joint neighborhood distribution of race and income.

**Estimation Approaches**

One approach to estimating $f^h_g(p, q)$ is to estimate $\rho_{gi}(p)$ in each neighborhood $i$. Then $R_{gi}(p)$ can be estimated as the integral of $\int_0^r \hat{\rho}_{gi}(r)dr$, and, because we can observe the group counts within each neighborhood, we estimate $f^h_g(p, q)$ as

$$f^h_g(p, q) = \sum_i \left[ \frac{T_i \cdot \pi_{gi} \cdot \hat{\rho}_{gi}(p)}{T \cdot \pi_g \cdot \hat{\rho}_g(p)} \cdot \frac{T_i \cdot \pi_{hi} \cdot \hat{R}_{hi}(q)}{T_i} \right].$$

The potential drawback of this approach is that it requires us to estimate $\rho_{gi}$ from small samples in each neighborhood $i$ and group $g$. Instead, we adopt an alternative approach.

Rather than directly estimating $f^h_g(p, q)$, we estimate instead $\hat{f}^h_g(p, q)$. From Equation (11), we can write $f^h_g(p, q)$ as
\[ f^h_g(p, q) = \frac{d}{dp} \left[ \frac{R_g(p) \cdot F^h_g(p, q)}{\rho_g(p)} \right]. \]  

(14)

Estimating \( f^h_g(p, q) \) therefore becomes a matter of estimating the functions \( R_g(p), \rho_g(p), \) and \( F^h_g(p, q) \). If we have race-by-income category-by-neighborhood household counts, we can compute \( R_{gi}(p) \) and \( F^h_g(p, q) \) for specific values of \( p \) and \( q \) (the values of \( p \) and \( q \) that correspond to the cumulative population proportions in or below each income category). \( R_{gi}(p) \) is simply the proportion of members of group \( g \) in neighborhood \( i \) with incomes in or below the income category whose upper bound is \( p \) in the population. We can then compute the exposure of members of group \( g \) with income less than or equal to \( p \) to members of group \( h \) with incomes less than or equal to \( q \):

\[ F^h_g(p, q) = \sum_i \left[ \frac{T_i \cdot \pi_{gi} \cdot R_{gi}(p)}{T^i \cdot \pi_g \cdot R_g(p)} \cdot \frac{T_i \cdot \pi_{hi} \cdot R_{hi}(q)}{T_i} \right]. \]  

(15)

In the 2007-2011 American Community Survey (ACS), for example, income is reported in 16 categories (e.g., “less than $5,000”; “$5,000-$10,000”; and so on). Suppose 5% of all households report incomes below $5,000 and 10% of all households report incomes below $10,000. Then we can compute \( F^h_g(.05,.10) \) and \( F^h_g(.10,.05) \) from equation (15) because we have counts in each neighborhood (tract) of the number of members of each race group with incomes below .05 and below .10.

Given the 16 income categories reported in the ACS, and assuming we divide the population into five mutually exclusive and exhaustive groups (non-Hispanic white, non-Hispanic black, non-Hispanic Asian/Pacific Islander, non-Hispanic other race, and Hispanic), we can compute \( F^h_g(p, q) \) for \( 16 \cdot 16 \cdot 5 \cdot 5 = 6,400 \) different combinations of \( g, h, p, \) and \( q \). We can then fit a multidimensional polynomial surface to these 6,400 observations to estimate the functions \( F^h_g(p, q) \).

In addition, we can compute \( R_g(p) \) for \( 16 \cdot 5 = 80 \) combinations of \( g \) and \( p \). We can then fit five
group-specific polynomials to these observations. We get \( \rho_g(p) = \frac{d}{dp} R_g(p) \). The Appendix describes the estimation of these functions in detail.

Data

In this paper, we use data from the 2007-2011 American Community Survey (ACS). We use census tracts as our definition of neighborhoods. The ACS provides partial cross-tabulations of household counts by income and racial/ethnic categories. In the 2007-11 ACS data, there are 16 categories of income, 7 race categories, and 1 indicator for whether the household is of Hispanic origin. We focus here on five mutually exclusive race/ethnic groups: non-Hispanic Asian, non-Hispanic black, Hispanic, non-Hispanic white, and non-Hispanic other. We use an iterative proportional fitting (IPF) process to estimate the full cross-tabulations of these five race/ethnic categories by income within each census tract, using Public Use Microdata Samples (PUMS) data to seed the IPF tables (for complete details on the construction of the cross tabulations and a discussion of the accuracy of the IPF process, see Reardon, Fox, & Townsend, forthcoming). The result is a dataset containing counts of the number of households in each race/ethnic group that are in each income category for each tract in the U.S. in 2007-11.

Illustrative Application of the Approach

Using the 2007-2011 ACS data and the estimation methods described in the appendix, we compute 80 observed values of \( R_g(p) \) and 6,400 values of \( F_g^h(p, q) \) for each of the values of \( g, h, p, \) and \( q \) observed in the ACS data. We fit multidimensional polynomials to these data to estimate the continuous functions \( R_g(p) \) and 6,400 values of \( F_g^h(p, q) \). From these, we compute the functions \( f_g^h(p, q) \) and the other derived functions described above. Using these, we construct a set of illustrative figures to demonstrate a number of ways that the estimates can be used to describe the joint neighborhood distributions of race and income.
A first step in estimating the exposure measures is calculating race-specific household income densities, described by the functions $\rho_g(p)$. These density functions are presented (stacked) in figure 1, which shows the proportion of households of a given race/ethnicity at each percentile of the national household income distribution. The horizontal axis is household income in percentiles and dollars from the national household income distribution. Reading the figure horizontally describes the income distribution within a race. The vertical axis is population proportion. Reading the figure vertically, then, describes the proportion of households of each race among all households at a given income percentile.

Figure 1 illustrates the unequal income distributions among white, black, Hispanic, and Asian households in the United States. Black households are disproportionately concentrated at the lower end of the income distribution; Hispanics are disproportionately in the bottom half of the distribution, while White and Asian households are disproportionately above the national median income. Nonetheless, a majority of low-income households are white in the United States, by virtue of their much larger population share.

Figure 2 presents the 25th, 50th, and 75th percentiles of the average neighborhood income distributions of households at each point in the income distribution. These income distributions are described by the function $f_t^z(p, q)$. Rather than plot the full surface described by $f_t^z(p, q)$, however, Figure 2 plots selected percentiles of neighborhood income distributions. To compute these values, we construct the function $f_t^{z-1}(p, c)$ (by numerically inverting $f_t^z(p, q)$) for the values $c \in \{25, 50, 75\}$ and for $p \in \{1, 2, ..., 99\}$. For example, to identify the 50th percentile income in the typical neighborhood of a household with 5th percentile income, we set $f_t^5(5, q)$ equal to 50 and solve for $q$ via numerical
interpolation. The horizontal axis represents a household's own income; the vertical axis represents neighborhood household income.

As an example of how to read this figure, consider households with income of $20,000, which is approximately at the 18th percentile of the national income distribution. Such households, on average, live in neighborhoods where the 25th percentile of the neighborhood household income distribution corresponds to the 20th percentile of the national household income distribution (about $22,000). The median of the average neighborhood household income distribution is roughly equal to the 42nd percentile of the national household income distribution (about $44,000). And the 75th percentile of the average neighborhood household income distribution is at about the 68th percentile of the national household income distribution (roughly $77,000). Although Figure 2 shows only the 25th, 50th, and 75th percentiles of the average neighborhood income distributions, recall one could similarly construct these lines for any desired percentile. Figure 1 makes clear that households with higher incomes live, on average, in neighborhoods with higher household income distributions.

The steepness of the lines in Figure 2 captures the association between a household's own income and the 25th, 50th, and 75th percentile household income in the neighborhood. As noted elsewhere, the steepness of these lines provide an intuitive measure of income segregation (Reardon et al., forthcoming). Consider the 50th percentile line. A flat line would indicate no or little income segregation: average households at any income level live in neighborhoods with the same average income distributions. Steep lines would indicate a strong association between one's own income and that of one's neighbors. As both axes are presented in percentile terms, the maximum value of the slope, averaged over the range of percentiles, is one. The lines are steeper in the right side of Figure 2, indicating that segregation among upper-income households is moderately larger than among lower-income households, consistent with other research (Bischoff & Reardon, 2014; Reardon & Bischoff, 2011; Reardon et al., forthcoming). The lines here have average slopes of roughly 0.25-0.35, suggesting
segregation is roughly one-quarter to one-third as high as its theoretical maximum (which would only occur if all households lived in neighborhoods where they and their neighbors had identical incomes.

While the exposure measures themselves are calculated in percentiles, the axes need not be presented in percentiles. The axes can be re-scaled and shown in dollars, or even log dollars. The dollar figures here, as well as in all following figures, are 2008 dollars and correspond directly to the thresholds of the 16 income categories in the ACS data. For convenience, here and elsewhere, the axes are labeled in terms of both income percentiles and dollars.

**Figure 3. Average Cumulative Neighborhood Income Distributions, by Own Household Income, 2007-2011**

Figure 3 presents similar information as Figure 2, but in a different way. Figure 3 presents $f^i_t(p, q)$ for values of $p \in \{5, 25, 50, 75, 95\}$ and for value of $q \in [0, 100]$. In Figure 3, the estimated income exposure function (the exposure of households with incomes of $p$ to those with incomes less than or equal to $q$) is drawn as a function of $q$ for various values of $p$. To see the connection between the two figures, note that the 50th percentile (red) line in figure 2 corresponds to where each of the lines in figure 3 crosses the value 50 on the vertical axis. The gray line in figure 3, representing households with income at the 5th percentile, crosses this line around where $q$ equals 38. This means that for households at the 5th percentile, half of their neighbors have incomes below the 38th percentile, and half have incomes above the 38th percentile, making the 38th percentile the median neighborhood income for households at the 5th percentile. Figure 2 shows this as well: on the red line, when $p$ equals 5, the 50th percentile neighborhood income is 38. Drawing the surface in this way more clearly shows how spread out the lines are as neighbor income increases. For example, again consider where each of the lines crosses the exposure equals 50 line. As $p$ increases, the space between the lines becomes increasingly large, indicating larger discrepancies in neighborhood median income as household income increases.
Figure 4. Average Neighborhood Racial Composition, by Household Income, All Households in U.S., 2007-2011

Figure 4 shows the average neighborhood racial composition for households of different incomes. These are given by the functions \( f^h_t(p, 1) \). In contrast to figure 1, Figure 4 shows the average racial composition of households’ neighborhoods, not the population proportions. On the vertical axis, the typical racial composition of the neighborhood sums to 100%, and the figure shows how the racial composition of the neighborhood changes as a function of own household income. Higher-income households, on average, have more white and Asian neighbors and fewer black and Hispanic neighbors than lower-income households.

Note that figure 4 looks relatively similar to figure 1. If neighborhoods were sorted perfectly by income, then these two figures would be identical, because every household would have only neighbors with their same income, who would, by definition, have the same racial composition as the population at that income level. That said, there are other patterns that may make Figure 4 similar to Figure 1. For example, if neighborhoods were sorted perfectly by race, but not at all by income within racial groups, Figures 1 and 4 would again be identical. Thus, the similarity of Figures 1 and 4 is not particularly informative about the relative extent of racial and income segregation that underlie them.

Figure 5. Average Neighborhood Racial Composition, by Household Income and Race, 2007-2011

Figure 5 is similar to Figure 4, but presents average neighborhood racial composition as a function of both householders’ race and income, as described by the functions \( f^{gh}_t(p, 1) \). For example, the top left panel shows the average neighborhood racial composition for Asian households only, conditional on their household income percentile. This panel shows, for example, that Asian households at the 50th percentile of the income distribution live in neighborhoods where, on average, roughly 50% of households are white, 10% are black, 20% are Hispanic, and 20% are Asian. Note that here, and
throughout the paper, income percentiles are always measured in terms of the overall national income distribution, not group-specific income distributions. Of course, the axes could be scaled to reflect race group-specific household income distributions, in percentiles or dollars, if that were the goal of the description.

One striking feature of figure 5 is the high proportion of same-race households in the neighborhoods of each race group across, regardless of income. For example, the average neighborhood racial composition for Asian households shows that, across the income distribution, nearly 20% of households in the neighborhood are Asian households, despite the fact that Asian households make up only roughly 5% of the population. A similar pattern is seen for each race group. This pattern, when compared to Figure 4, indicates high levels of racial segregation, even conditional on income.

**Figure 6. Neighborhood Median Income, by Household Income and Race, All Households in U.S., 2007-2011**

Figure 6 describes the average neighborhood income distributions for households of different races, by household income. Each line describes the median income in of the average neighborhood income distribution for households of a given race group, as a function of their income. The lines come from the $f_{d}^{-1}(p,50)$ functions. The figure is similar to Figure 2 (which shoes the $f_{i}^{-1}(p,50)$ function), but shows only the median of the average neighborhood income distribution (not the 25th and 75th percentiles), and presents a separate line for each race group.

The most striking feature of Figure 6 is that Asian and white households live in neighborhoods where the median income of their neighbors is much higher than in the neighborhoods of similar income Hispanic and black households. The vertical distance between the lines yields a comparison of neighborhood conditions between households of different races. As in figure 2, the steepness of the lines is an indication of the degree of income segregation within each race group. The figure indicates that income segregation is higher for all groups in the upper half of the income distribution relative to the
lower half of the income distribution. Another way to compare neighborhood conditions of the different groups is to look at the horizontal differences between the lines. From this perspective, the figure shows that black and Hispanic households typically live in similar neighborhoods (in terms of their median income) as white households with much lower incomes. Black households with incomes of roughly $60,000, for example, live in neighborhoods similar to those of white households earning roughly $12,000 (see Reardon et al., forthcoming).

**Figure 7. Race-specific Neighborhood Median Income, by Household Income, All Households in U.S., 2007-2011**

Figure 6 above illustrates the income distribution in the typical neighborhood of households of different races. Figure 7, in contrast, illustrates the opposite: the income distributions each race in the average neighborhood of households of a given income. Specifically, Figure 7 plots the functions \( f_{t_{h^{-1}}}(p, 50) \). Each line represents the median of the income distribution of a specific group for typical households of specific incomes. For example, a typical household at the 25th percentile of the income distribution lives in a neighborhood where the median black income is at roughly the 30th percentile of the national income distribution and the median white income is at roughly the 47th percentile. The figure shows that, on average, Asian and white income distributions are higher than those of blacks and Hispanics for all values of household income, suggesting that, for most households in the United States, their black and Hispanic neighbors are poorer than their Asian and white neighbors.

**Figure 8. Race-specific Neighborhood Median Income, by Household Income and Race, 2007-2011**

Figure 8 is similar to Figure 7, but shows race-specific median incomes in the neighborhoods of households of each race. Specifically, each line in Figure 8 is one of the \( f_{g_{h^{-1}}}(p, 50) \) functions. For example, the green line in the top left panel indicates that the typical Asian household at the median of
the national income distribution lives in a neighborhood where the median income among Hispanic households is around the 46th percentile of the national income distribution. As in the other figures showing the $f^{-1}$ functions, one could choose other percentiles of these race-specific distributions to display as well.

In Figure 8, the steepness the lines indicates the degree of within-race segregation. The top left panel suggests, for example, that Asian households are highly segregated by income—low income Asians live in neighborhoods where their Asian neighborhoods are poor, on average, while high income Asians have much higher-income Asian neighbors. Within-group income segregation is also high for black households, but is lower among white and Hispanic households.

**Figure 9. Metropolitan Variation in Neighborhood Median Income, by Household Income, Ten Largest Metropolitan Areas by Population, 2007-2011**

Each of the previous figures describes patterns for the United States as a whole. The methods we have described here, can be applied to compare smaller geographic regions as well. Figure 9 provides an example of this. It shows median neighborhood income, as a function of household income, for each of the 10 largest metropolitan areas in the United States. The lines come from metropolitan area-specific functions $f_{mt}^{-1}(R_m(p),50)$, where $R_m(p)$ is a function that converts national income percentiles to local income percentiles of metropolitan area $m$ (that is, it is the cumulative income distribution function for metropolitan area $m$). The figure indicates, for example, that households at the 50th percentile of the national income distribution in the Washington, D.C. metropolitan area live in neighborhoods where the median income is above the 65th percentile of the national income distribution.

A notable feature of this figure is that both axes are shown in the national income distribution to allow for comparisons across metropolitan areas. Although the graph could be constructed using local income distributions, that would obscure comparisons among households of the same income in
different metropolitan areas, because the 50th percentile of Chicago’s income distribution is not the same as the 50th percentile of New York’s income distribution. Using a common scale for income (percentiles of the national income distribution) makes evident that households in some metropolitan areas live, on average, in very different neighborhoods than similar income households in other areas. For example, Washington, DC households earning $60,000 live in much higher income neighborhoods than do similar income households in Los Angeles; in fact, Washington, DC households earning $60,000 typically live in neighborhoods similar to those of Los Angeles households earning $150,000.

Discussion

The approach we have outlined here provides a variety of ways of characterizing the joint patterns of racial and socioeconomic segregation. A full characterization of these patterns is provided by the group-specific income distributions (the $R_g$ functions, in our notation) and the set of exposure functions that describe the average neighborhood income distributions conditional on race and income (the $f^{gh}_g$ functions in our notation), but simply reporting the parameters of these functions is neither feasible nor particularly informative (in our illustration here, these functions are characterized by a set of 420 parameters). Instead, we have chosen to illustrate key features of these functions in a series of figures, each of which highlights a different aspect of the joint distribution.

One could, of course, derive additional statistics from these functions. The slopes of the lines in Figures 2 and 8, for example, may be useful as measures of segregation. The vertical or horizontal distances between the lines in Figure 6, likewise, might be thought of as measures of racial inequality in neighborhood conditions net of differences due to between-race differences in household income. Measures of between-group differences in racial composition of neighborhoods (evident in Figure 5) may be useful for measuring and understanding racial segregation. Statistics of these types can be derived from the estimated $R_g$ and $f^{gh}_g$ functions and then may be usefully compared across time or metropolitan
areas to assess changes or variation in patterns of racial/economic segregation.

Our goal here was to describe a general approach to measuring joint patterns of racial and socioeconomic segregation. Given that, a discussion of the substantive implications of the patterns illustrated in our figures here is beyond the scope of this paper, but a few features of the figures are particularly striking. First, the figures clearly show that there are large racial differences in neighborhood racial and economic composition, even conditional on income. That is, equally poor white, black, Hispanic, and Asian households are located in very different neighborhoods than one another. This is very consistent with prior research showing that economic disparities are insufficient to explain racial segregation (Adelman, 2004; Logan, 2002, 2011; Reardon et al., forthcoming; Timberlake, 2002, 2007; Timberlake & Iceland, 2007). If racial segregation were simply the result of racial differences in income, we would expect racial differences in neighborhood composition to disappear once we condition on household income. The figures here clearly show that they do not.

Second, the figures reveal something about the income levels of households that different racial and income groups might encounter in their neighborhood. Figures 7 and 8 show that the typical household, regardless of income level or race, lives in a neighborhood where black and Hispanic neighbors have lower incomes than white and Asian neighbors. Indeed, the black and Hispanic neighbors of high-income households have lower median incomes, on average, than the white and Asian neighbors of low-income households (see Figure 7). This pattern may play a role in shaping racial stereotypes and perceptions.

Third, Figure 9 shows there is substantial variation among metropolitan areas in the patterns of exposure to high- and low-income neighbors, conditional on income. Not shown here, but straightforward to compute from the methods described here, are metropolitan patterns of racial differences in neighborhood economic conditions. A full description of variation across metropolitan areas in the joint neighborhood distribution of race and income would likely reveal considerable
variability.

If neighborhood economic conditions affect child development and opportunities for educational and economic success, as suggested by the neighborhood effects literature, then the variation in economic neighborhood conditions across racial groups, households of different incomes, and metropolitan areas may lead to disparities in developmental, educational, and economic outcomes. The patterns of racial disparities in neighborhood conditions evident in Figure 6, for example, help to quantify the extent to which black and Hispanic children are disadvantaged by neighborhood context above what can be accounted for by their incomes. And the patterns evident in Figure 9 suggest that growing up in a low-income household in Los Angeles may be much worse than growing up in an equally poor household in Washington, DC.

The methods described here provide a consistent way of quantifying disparities in neighborhood conditions that is largely independent of the specific income thresholds used in tabulating income in the ACS data. This makes possible much more precise comparisons of racial and economic neighborhood conditions across place, time, and population groups than has been used in prior work. We hope that these methods enable researchers to more carefully investigate the patterns, causes, and consequences of racial and economic segregation.
Figures

Figure 1. Race-Specific Income Distributions, All Households in U.S., 2007-2011

Note. Figure 1 presents the proportion of households in each race group at a given household income percentile for all households in the U.S. for the years 2007-2011. The x-axis is household income, on the national scale, in percentiles and dollar figures; the y-axis is proportion. The figure shows that, for example, there is a higher proportion of white households than all other groups combined across the distribution of household income, and that there are a higher proportion of white households in the upper end of the income distribution than the lower end.
Figure 2. Percentiles of Average Neighborhood Household Income Distributions, by Own Household Income, 2007-2011

Note. Figure 2 presents neighborhood 25th, 50th, and 75th percentile household income, conditional on own household income, for all households in the U.S. for the years 2007-2011. The x-axis indicates household income; the y-axis indicates median household income in the neighborhood of a typical household of a given income. For both axes, the percentiles and dollar figures are taken from the national household income distribution. As an example of how to read the table, consider households earning $60,000/year (roughly the 56th percentile of the household income distribution). Such households live, on average, in neighborhoods where the 25th percentile household income is $27,000 (roughly the 27th percentile of the national household income distribution), the 50th percentile household income is about $53,000 (roughly the 50th percentile of the national household income distribution), and the 75th percentile household income is about $90,000 (roughly the 72nd percentile of the national household income distribution).
Figure 3 presents the exposure of households with various incomes to the distribution of cumulative household income for all households in the U.S. for the years 2007-2011. The y-axis represents the exposure measure, which ranges from 0 (no exposure) to 100 (complete exposure). The x-axis represents the national cumulative household income distribution in both percentiles and selected dollar values. The lines, then, trace the exposure of those with incomes at the 5th, 25th, 50th, 75th, and 95th percentile household incomes to those with incomes equal to or less than the values presented on the x-axis. As an example of how to read this figure, consider the red line, and a y-axis value of 50. This value of exposure allows us to identify the median income in a given neighborhood – half of all households have incomes below this value, and half have incomes above this value. The red lines suggest that for households that are themselves at the 75th percentile on the household income distribution, the median income in their typical neighborhood is at approximately the 57th percentile of the national household income distribution.
Figure 4. Average Neighborhood Racial Composition, by Household Income, All Households in U.S., 2007-2011

Note. Figure 4 presents the average neighborhood racial composition by household income, for all households in the U.S. for the years 2007-2011. The x-axis is household income, on the national scale, in percentiles and dollars. The y-axis is proportion. As an example of how to read the figures, consider households at the 40th percentile of the household income distribution. The typical neighborhood household racial composition for households at this income level is close to 64% white, 15% black, 15% Hispanic, 5% Asian, and 1% other.
Figure 5. Average Neighborhood Racial Composition, by Household Income and Race, 2007-2011

Note. Figure 5 presents four panels, each of which shows the average neighborhood racial composition by household income, for households of a specific racial/ethnic group for the years 2007-2011. The x-axis is household income, on the national scale, in percentiles and dollars. The y-axis is proportion. Panel 1 presents neighborhood composition for Asian households, panel 2 for black households, panel 3 Hispanic households, and panel 4 white households. As an example of how to read the figure, consider households at the 40th percentile of the household income distribution in panel 1. The typical racial composition in the neighborhoods of Asian households at this income level is close to 50% white, 10% black, 18% Hispanic, 20% Asian, and 2% other.
Figure 6. Neighborhood Median Income, by Household Income and Race, All Households in U.S., 2007-2011

Note. Figure 6 shows neighborhood median household income, conditional on own household income and race/ethnicity, for all households in the U.S. for the years 2007-2011. The x-axis is own household income; the y-axis is neighborhood median household income. For both axes, the percentiles and dollar figures are taken from the national household income distribution. As an example of how to read the table, consider White households at the 50th percentile of the national household income distribution. The y-axis indicates that such families live, on average, in neighborhoods where the median income is close to $55,000, very close to the median of the national distribution.
Figure 7. Race-specific Neighborhood Median Income, by Household Income, All Households in U.S., 2007-2011

Note. Figure 7 shows race-specific neighborhood median household income, conditional on own household income, for all households in the U.S. for the years 2007-2011. The x-axis is own household income; the y-axis is neighborhood median household income. For both axes, the percentiles and dollar figures are taken from the national household income distribution. As an example of how to read the table, consider households at the 50th percentile of the national income distribution (the value of the x-axis). The y-axis indicates that such families live, on average, in neighborhoods where the median income of blacks is around the 36th percentile of the national income distribution and the median income of whites is around the 58th percentile.
Figure 8. Race-specific Neighborhood Median Income, by Household Income and Race, 2007-2011

Note. Figure 8 shows race-specific neighborhood median household income, conditional on own household income and race/ethnicity, for the years 2007-2011. The x-axis is own household income; the y-axis is neighborhood median household income. For both axes, the percentiles and dollar figures are taken from the national household income distribution. As an example of how to read the table, consider White households (bottom right) at the 50th percentile of the national household income distribution (the value of the x-axis). The y-axis indicates that such families live, on average, in neighborhoods where the median income of blacks is around the 37th percentile of the national income distribution and the median income of whites is around the 60th percentile.
Figure 9. Metropolitan Variation in Neighborhood Median Income, by Household Income, Ten Largest Metropolitan Areas by Population, 2007-2011

Note. Figure 9 presents neighborhood median household income, conditional on own household income, by metropolitan area for the years 2007-2011. The $x$-axis indicates household income; the $y$-axis indicates median household income in the neighborhood of a typical household of a given income. For both axes, the percentiles and dollar figures are taken from the national household income distribution (not from each metropolitan area). The markers on the lines indicate the 25th, 50th, and 75th percentiles of each metropolitan area’s own household income distribution. As an example of how to read the figure, consider households in Minneapolis-St. Paul Bloomington, MN-WI at the 60th percentile of the national income distribution (roughly $66,000). These households typically live in neighborhoods of the Minneapolis-St Paul metropolitan area with median incomes of roughly $64,000, about the 59th percentile of the national income distribution.
References


segregation in cities (pp. 61-82). London: Croon Helm.


