

Birds of a Feather: Social Bases of Neighborhood Formation in Newark, New Jersey, 1880

John R. Logan¹ · Hyoung-jin Shin²

Published online: 6 July 2016
© Population Association of America 2016

Abstract This study examines the bases of residential segregation in a late nineteenth century American city, recognizing the strong tendency toward homophily within neighborhoods. Our primary question is how ethnicity, social class, nativity, and family composition affect where people live. Segregation is usually studied one dimension at a time, but these social differences are interrelated, and thus a multivariate approach is needed to understand their effects. We find that ethnicity is the main basis of local residential sorting, while occupational standing and, to a lesser degree, family life cycle and nativity also are significant. A second concern is the geographic scale of neighborhoods: in this study, the geographic area within which the characteristics of potential neighbors matter in locational outcomes of individuals. Studies of segregation typically use a single spatial scale, often one determined by the availability of administrative data. We take advantage of a unique data set containing the address and geo-referenced location of every resident. We conclude that it is the most local scale that offers the best prediction of people's similarity to their neighbors. Adding information at larger scales minimally improves prediction of the person's location. The 1880 neighborhoods of Newark, New Jersey, were formed as individuals located themselves among similar neighbors on a single street segment.

Keywords Ethnic neighborhood · Residential segregation · Discrete choice model

Introduction

Neighborhood patterns of immigrants and minorities have drawn much attention from social scientists because racial and ethnic group boundaries are manifested through

✉ John R. Logan
john_logan@brown.edu

¹ Department of Sociology, Brown University, Box 1916, Providence, RI 02912, USA

² Department of Sociology, Kyungpook National University, 1370 Sankyuk-dong, Puk-gu, Daegu 702-701, Korea

residential segregation (Lieberson 1963). People are also segregated by social class, partly through self-selection but primarily by the cost of housing (Jargowsky 1996), and many studies of social ecology (following Shevky and Bell 1961) have documented separation on the basis of family life cycle. We investigate the neighborhood patterns of residents of Newark, New Jersey, in 1880. The key theoretical question is how ethnicity, class, and other social characteristics organize neighboring, where “neighboring” is defined not by social interaction but instead by living in the same local area. Although the data are from more than a century ago, this question remains important today, and our approach to dealing with it is relevant to any period. We apply discrete choice models to analyze how individual characteristics (such as ethnicity and occupation) are related to place characteristics (such as the percentage coethnic neighbors and average occupational level). Using these models, we draw conclusions by comparing the characteristics of the place where people actually live with all other places in the same city where they might have lived. No previous research has examined neighborhoods in this way in a historical setting.

Sorting People Into Neighborhoods

Our subject is the neighborhood and, specifically, how neighborhoods are constituted by the sorting of residents into different areas. Spatial differentiation was a central observation of early human ecologists who noted “the continuous sifting and sorting of the city’s population...[which produces] a patchwork of local areas differentiated from one another by cultural, racial, or linguistic peculiarities” (McKenzie 1921/1968:73; see also Burgess 1925:56). Contemporary social scientists agree that the spatial organization of cities is multidimensional. A neighborhood is a “bundle of spatially based attributes” (Galster 2001:2012). Seeking to simplify this bundle, Shevky and Bell (1961) argued that local areas within the city could be delineated on the basis of indicators of just three dimensions: social rank, urbanization, and segregation. A considerable literature in factorial ecology has pursued this approach by applying factor analysis to summarize the correlations among large numbers of census tract characteristics (Berry and Rees 1969; Van Arsdol et al. 1958; White 1987). Such studies are based on aggregate data. Indicators with high ecological correlations establish the dimensions of neighborhood differentiation, and these can then be used to define clusters of similar areas as neighborhoods (Sampson et al. 1997).

We do not seek to identify neighborhoods but rather to understand the processes through which they are created and maintained by individuals’ choices about where to live. This purpose requires simultaneous information at the aggregate level (e.g., the composition of small areas) and at the individual level (e.g., the set of characteristics of people who live in these areas). This multilevel approach is similar in spirit to the locational attainment models advanced in a series of studies by Alba and Logan (1992; see also Logan et al. 2002). A disadvantage of those models is that they examine one neighborhood outcome at a time. In practice, people are faced with potential neighborhoods that vary simultaneously on many dimensions. To address this concern, we employ discrete choice models of the sort estimated by Bruch (2014), in which the specific neighborhood in which a person lives is predicted by the combination of local area and personal characteristics (see also Bruch and Mare 2012; Quillian 2015). As

Quillian (2015:258) pointed out, “The major advantage of the discrete-choice approach is its ability to represent place outcomes as multidimensional...or as composed of a bundle of attributes that simultaneously influence locational attainment. Discrete choice then allows us to address the question of the relative importance of each neighborhood attribute in determining destinations.”

Here, we briefly review the literature on which social characteristics are likely bases of neighborhood formation. Most attention has been given to race/ethnicity and social class, and how these two dimensions interact. There are several points of view. Although this study cannot test the hypothesized processes, they are important motivators for this research.

Spatial Assimilation

In the spatial assimilation tradition, it is assumed that ethnic residential segregation will be undermined over time by socioeconomic mobility and acculturation of immigrant group members (Massey 1985). In other words, segregation by ethnicity will give way to more powerful forces related to class and nativity.

Place Stratification

In contrast, the place stratification perspective argues that some racialized groups will experience segregation regardless of their economic resources. Studies of contemporary residential segregation have documented that blacks in the United States are less likely to convert their socioeconomic resources into better residential outcomes than other groups because they are more susceptible to prejudice and discrimination (Charles 2003; Massey and Denton 1993; White 1987). Many empirical studies have shown that the effect of class standing on residential outcomes is relatively smaller for blacks than for Hispanics or Asians (Alba et al. 2000; Denton and Massey 1988; Iceland and Wilkes 2006; Logan et al. 2004). These processes are susceptible to change over time. For example, Wilson (1987) noted that middle class blacks in the 1940s and 1950s were required to live in urban black neighborhoods (race trumped class) but increasingly were able to move out of such neighborhoods by the 1980s (class began to play a larger role in their locational outcomes).

Organization of Production

Another factor affecting residential outcomes is the proximity of workplaces. Scott (1990) made the general argument that the organization of production affects the spatial distribution of occupations in a city. In the late nineteenth century context that we study, this effect could show up in both ethnic and class segregation, but its force depended on transportation infrastructure. While intercity and intracity rail transport existed, the average city dweller still likely walked everywhere (Hershberg 1981). As new transit systems lowered the economic barriers to commuting, increasing numbers of Americans left crowded city centers for more peripheral neighborhoods, thereby increasing geographic separation of rich and poor (Warner 1962). At the extremes, the capability of the middle and upper classes to separate home and work by commuting into the inner city supported increasing class segregation.

Preferences

A fourth theoretical approach gives priority to preferences. Living together with people of similar social class may be due not only to comparable financial resources but also to greater satisfaction from living among peers. Logan and colleagues (2002) argued that some ethnic groups have such strong preferences for ethnic communities that immigrants with more financial resources and human capital may choose to live in ethnic neighborhoods even though they have an option to enter the mainstream community. Similar preferences could lead people to select neighborhoods based on their family composition (e.g., unmarried persons may prefer different areas than families with children) and nativity (aside from coethnic preferences and economic constraints, immigrants may prefer neighborhoods with more immigrants).

Little systematic evidence is available on these alternative theories for the late nineteenth century. An important exception is Zunz's (1982) study of Detroit in the years 1880–1920. He argued that segregation here was initially (in 1880) primarily based on ethnicity. He found that by 1920, Poles, Hungarians, Jews, and blacks “reached record levels of concentration in some blocks” (1982:341). During this same period, however, Irish concentrations disappeared, and German clusters became more scattered. Simultaneously, occupation became more important to location, often in tandem with ethnicity. In 1880, ethnic blocks commonly included a wide range of occupations, from laborers to professionals and shopkeepers. By 1920, however, many whole city blocks were made up entirely of people with the same ethnicity and occupation, such as Polish factory workers, native white factory workers, or native white office workers.

Other historians have linked Newark's class and ethnic structure to the changing relationship among jobs, ethnic niches, and neighborhood. Newark's original business district on the western bank of the Passaic River included the city's most desirable residential property in the mid nineteenth century. “The convenience of living close to the hub of economic, social and religious activities, the security offered by good police protection, and the inadequate facilities for intra-city travel, combined to keep residential real estate in the inner wards expensive and exclusive” (Popper 1952:160–161). Yet, by 1870, seven horse car routes extended even beyond the city limits, facilitating early suburbanization. At the same time, new railroad freight lines were introduced, with terminals, warehouses, and industrial plants located near the tracks (Drummond 1979:115, 131). Consequently, Newark's neighborhoods were being divided in new ways. Many of Newark's Irish immigrants worked in the docks along the Passaic and lived in poor areas nearby, while many Germans worked on farms in the less-populated western edge of the city. Class, ethnicity, and neighborhood were clearly interlocked.

The Spatial Scale of Neighborhoods

A second goal of this study is to discover the spatial scale at which social homophily organizes neighborhoods. The common practice in ecological analysis is to identify neighborhoods by administrative units for which relevant data are available, such as census tracts or combinations of census tracts—what White (1987) referred to as “statistical areas”—despite widespread agreement that such units are arbitrary (Dietz

2002). Although this practice has been imposed by the lack of data at smaller scales, it has a theoretical justification. A “neighborhood” of as many as 3,000–5,000 residents (the usual range for a contemporary census tract) is large enough to constitute a market for local goods and services, and may support institutions, such as a school or church. However, other spatial scales also have substantive meaning. As Reardon et al. (2008:490, citing Kaplan and Holloway 2001) pointed out, “There is no single geographic scale of segregation.”

Urban sociologists are familiar with the coexistence of different vertical layers of local social organization, often starting with the *face block* (both sides of a street between two intersections) extending to a larger neighborhood, community area or district, or even an entire city (Hunter 1974; Suttles 1972). Both Suttles and Hunter considered the *street segment* (a residential face block bounded by intersecting streets) or face block to be the main basis of casual social relationships and face-to-face neighbor relations. Anderson (1992) and Grannis (2009) made the same assumption. Suttles (1972:56) referred to the next layer (“the smallest area which possesses a corporate identity known to both its members and outsiders”) as the *defended neighborhood*—the area in which he expected to find collective mobilization. Kusenbach (2008) more recently investigated neighborhoods at various spatial scales (microsettings within blocks, street blocks, walking-distance neighborhoods, and larger enclaves), showing that they have different constellations of residents’ sentiments and practical uses of their environment, neighborly interaction and relationships, and locals’ participation in collective events and rituals.

Studies using agent-based models of residential location are also confronted with the question of spatial scale. O’Sullivan (2005) presumed that people make choices about where to live (or whether to move) based on the composition of their neighborhood at two scales. One is the *local neighborhood*, comprising the immediate neighbors who live adjacent to the person. This is the scale of what Schelling (1978:147) termed a “self-forming neighborhood” because in his simulations, even moderate levels of intolerance for difference can lead to a socially homogeneous area at this small scale. The other is a larger *bounded neighborhood* containing many households, such as a school attendance area or an area with a distinctive locational reputation. When characteristics of bounded neighborhoods are important to agents (e.g., because they value sending children to a particular school), O’Sullivan (2005) believed that segregation at this larger scale may be greater than at the local scale. Hence, neighborhoods can be organized at either or both of these spatial scales.

Because the agent-based model conceptualizes neighborhoods as the environment surrounding a given person, it questions a key assumption of ecological studies: namely, the assumptions that neighborhoods actually are clearly bounded and that they do not overlap. That is, at some borderline, one neighborhood ends, and another begins. The alternative is to study what geographers call “egocentric neighborhoods.” That is, every person’s residential or work location is the center of their own neighborhood, so that neighborhoods are inherently overlapping (Chaix 2009; Spielman and Logan 2013). The idea that neighborhoods exist around people is reinforced by studies of people’s perceptions. Field experiments by Coulton et al. (2001) found that most residents placed their own home at the center of their neighborhood. Survey research by Hunter (1974) reported that neighborhoods had *rolling boundaries*: that is, people might agree on the name of their neighborhood, but those living near its edge tend to

perceive it as extending further in that direction. Egocentric neighborhoods can also be constructed at multiple scales. For example, Lee et al. (2008) used a series of concentric rings to estimate spatial measures of residential segregation at different distances from a person's home. We do not dispute the value of thinking about neighborhoods as fixed, bounded, and nonoverlapping areas of the city. However, the advantage of studying egocentric neighborhoods is that we can evaluate more precisely the social composition of the areas surrounding a person without making *a priori* assumptions about geographic scale and without having to treat people who live at the edges of neighborhoods as living in the same locale as people who are located at the neighborhood core.

Research Design

This study uses data from the city of Newark, New Jersey, in 1880. It relies on 100 % individual-level microdata from the 1880 census and the ability to identify the specific location of each person's residence in the city. Information about characteristics of individuals is compared with information about neighborhoods at multiple spatial scales through discrete choice models. This approach is unique in two crucial ways. First, most previous historical research on segregation (see the sources on black-white segregation cited by Massey and Denton (1993) and Lieberman's (1963) groundbreaking studies of white ethnic segregation) relied on data after 1900 and aggregated at the level of city wards. Wards can be very large areas (Chicago's wards in 1900, for example, averaged a population of about 50,000) and are thus not well suited to identifying neighborhoods. We are able to define neighborhoods at any spatial scale. Second, previous historical studies have been limited in their ability to assess segregation on dimensions other than race and ethnicity. We will draw systematically on race/ethnicity, nativity (foreign birth), occupational standing, and family composition.

Newark is one of 39 cities for which geocoded full count microdata are available from the Urban Transition HGIS Project (Logan et al. 2011). We selected Newark for this analysis because it is among the few cities with large shares of immigrants from both Ireland (typical of Eastern Seaboard cities) and Germany (typical of Midwestern cities). Irish immigrants flocked to Newark in the 1820s to work on the construction of the Morris Canal. Construction jobs at the canal were dangerous and underpaid, but the Irish, who were mostly poor and unskilled laborers, accepted these conditions (Wepman 2002). Later, the Irish Great Famine during the 1840s stimulated a larger wave of immigrants, who were met with rising anti-Irish sentiment (Bider and Reimers 1995). Germans also began arriving in Newark in significant numbers during the 1840s. While the Irish were often described as an "unassimilable" population because of their low socioeconomic status (SES) and Catholic religion, Germans were portrayed as affluent farmers or skilled artisans but standing culturally apart from local people. Future studies can extend this analysis to a comparison across cities. It could be important, for example, to ask whether ethnicity is more consequential for neighborhood formation in cities with a single dominant ethnic minority or in one where many are represented.

Data

Utilizing the 1880 full-count census data prepared by the Minnesota Population Center, we developed a GIS street map of Newark, NJ, and geocoded individual addresses. This study uses a data set in which 133,554 persons (98 % of Newark's total population) in 28,489 households are georeferenced. The 100 % sample for the 1880 census includes information about several key population characteristics. These individual-level variables can be aggregated to local areas at any spatial scale for all the focal person's neighbors but not including data for that person.

In the descriptive analyses (characteristics of the city population and levels of segregation and isolation), the calculations are based on persons 18 and older (80,116 men and women). This age cutoff decision reflects that location decisions are not made by children (although parents may take them into account in choosing locations) and also avoids an ambiguity about whom to treat as a native white person. Almost all children living with their second-generation parents were born in the United States; on that basis, they could be considered native whites (3+ generation). However, we believe that such classification would imply greater ethnic diversity than actually existed. Children are therefore excluded from all of our analyses.

In multivariate analyses of who lives where, we make further selections. The finest geography that we study here is the street segment. Because we wish to take into account neighbors' SES as a neighborhood characteristic, we select only street segments (and residents of street segments) where there are enough cases (five or more neighbors with a listed occupation) to construct a reasonably reliable measure. This selection reduces the sample from 1,498 to 1,442 street segments. Because the majority of women did not have a listed occupation—and to avoid including multiple family members in the multivariate analysis (e.g., husbands and wives or fathers and coresiding adult daughters)—the analysis is limited to men (age 18 and older) who had a job in 1880. In most cases, this means that we select only one person from each household. These procedures reduce our individual-level sample to 28,922 persons. Following established practice (Duncombe et al. 2001; Mare and Bruch 2003; McFadden 1978), we impose a further limit on the sample size (the number of individuals and the number of street segments on which they resided) in order to stay within our computational capacity for discrete choice models (the computational load is the number of individuals multiplied by the number of street segments, more than 4 million observations.) After some experimentation, we found that the optimum size was a 10 % random sample (2,894 individuals), including as possible options all 1,442 residential street segments in which a sampled person of any group lived.

We include all residents, recent movers, and stayers in this sample. We do this mainly for substantive reasons: as Bruch and Mare (2012:123) pointed out, "Nonmoves are informative about residential choice because it is likely that the chances of opting for one's own neighborhood do in fact depend on the measured characteristics of the neighborhood." This is the choice made by Bruch and Mare (2012) and Bruch (2014). It is also imposed by the 1880 census data, which do not report when people moved to their current residence. In contrast, Mare and Bruch (2003) limited their analysis to recent movers. This decision makes model estimation more manageable, and it also has a substantive rationale. The neighborhood today may be very different from the neighborhood that previous movers came to years ago, and those previous movers

themselves may have changed over time. The temporal link between personal and neighborhood characteristics is stronger for recent movers. We suspect that our results are heavily influenced by choices made by recent movers. First, more than 40 % of the population was foreign-born. Second, the city had grown by nearly 30 % between 1870 and 1880, and many of the inhabited street segments in 1880 had not yet been developed a decade earlier. Third, renters in tenement buildings made up a large share of the population, and they were known to move frequently. For example, Erickson (1995) described the culture of “moving day” in nearby New York City, where it was routine for tenants to move to a new rental at the beginning of May when their leases expired.

We focus on four key individual characteristics: ethnicity/race, nativity, SES, and family status. We use the person’s and parents’ race and place of birth to create categories of ethnicity/race and nativity. We treat as native white (more than one-quarter of the population) those persons who were born in the United States and whose parents are also native-born white (usually referred to in the literature as native white of native parents). For the foreign-born, their country of birth determines their ethnicity. For those who were born in the United States but for whom at least one parent was born abroad, ethnicity is primarily determined by the mother’s country of birth. If only the father was foreign-born (or if the mother was foreign-born but her birthplace was not reported), the father’s country of birth is applied. The main categories are Irish, German, and British. A much smaller share of the population was black, including those categorized as black or mulatto regardless of the person’s and parents’ place of birth. A residual “other” category includes diverse European nationalities and a very small share of nonwhites.

Nativity is measured by country of birth: native (born in the United States) and foreign (born elsewhere). As a measure of social class, we use occupational socioeconomic index (SEI), an indicator originally calibrated to mid-twentieth century data but shown to be a valid indicator of relative position in this period (Sobek 1996). When SEI is treated as a categorical variable to measure segregation, we distinguish categories of high (more than 30), middle (20–30), and low (less than 20). Family life cycle is indicated by a dummy variable identifying persons who are married and living with at least one own child versus all others.

These individual characteristics are introduced in the discrete choice model as interaction terms involving place characteristics. We compute these place variables from the whole population of Newark, not from the individual sample: (1) the mean occupational SEI of the neighborhood; (2) the percentage foreign-born; (3) the percentage of married persons with a child; and (4) the percentage of each major racial/ethnic group. With access to the full microdata, we can measure these place characteristics without including the focal person in our sample, which is particularly helpful when neighborhoods are measured at a fine scale. Especially for our measure of occupational standing, many street segments have very few employed persons, and including the focal person in the place measure would create an artificial correlation between the person and neighborhood characteristic.

We measure these contextual variables at three scales of neighborhood. At the finest level, we aggregate various individual characteristics (but not including the sampled person) from a single street segment. We also study location in a larger set of connected street segments that we call a “segment group.” The relationship between a street

segment and segment group is illustrated in Fig. 1. A segment group is the larger neighborhood area within which a street segment is embedded. Finally, we add another layer of street segments that are connected with those in the segment group. We call this the “extended segment group.”

Note that neighborhoods at all three scales are egocentric: they are formed from the perspective of the resident of a given street segment. This means that for all 1,442 street segments, the number of segment groups and extended segment groups is the same, and we can link the personal characteristics of every sampled person to those of other residents of the segment, segment group, and extended segment group. Our research question on spatial scale is how much people’s residential location depends on the characteristics of the specific street segment where they live, and how much better the prediction is given when we take into account information about the surrounding areas.

Discrete Choice Models

The multivariate analysis is based on discrete choice models for residential location (Duncombe et al. 2001; Bruch and Mare 2006, 2012; McFadden 1973, 1978; Quillian 2015). The discrete choice model is suited for conditions where people choose from among a set of options. It can address the questions of (1) what neighborhood characteristics affect the residential choice of people and (2) how their individual characteristics interact with neighborhood contexts. Its unique feature is that the characteristics of the place where people actually live can be compared with all other

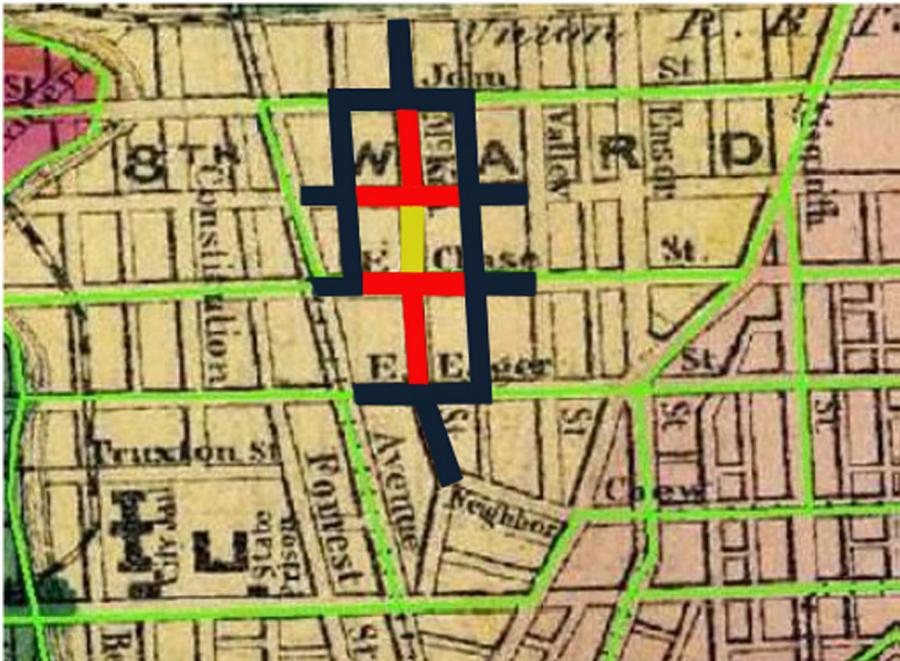


Fig. 1 Illustrations of three spatial scales: focal street segment (yellow), first-order neighbors (connected segments, red), and second-order neighbors (connected segment groups, black). Enumeration district boundaries are shown in green

places where they might have lived. Individual characteristics are introduced in the discrete choice model as interaction terms that are matched to the corresponding neighborhood characteristics. The effects of these interaction terms can be understood as reflecting spatial sorting mechanisms of homophily in residential choice. For example, statically significant and positive interactions between individual SEI and the mean SEI of the neighborhood would indicate that persons with a higher SEI are more likely to live in a neighborhood with a higher mean SEI. Ultimately, we will test the extent to which observed group differences in residential patterns are attributable to four different social dimensions (i.e., nativity, race/ethnicity, SEI, and family status).

Discrete choice models are usually expressed as conditional logit models in which the choice probabilities are the exponentiated value of log-odds coefficient estimates (Hoffman and Duncan 1988). The baseline model for discrete choice of residential location can be expressed as

$$P_{ij} = \exp(\mathbf{Z}_{ij}\alpha) / \sum_{k=1}^j \exp(\mathbf{Z}_{ik}\alpha). \quad (1)$$

In the baseline model, P_{ij} is the probability that individual i lives in neighborhood j . \mathbf{Z}_{ij} stands for the characteristics of the j th neighborhood for individual i , and α denotes the vector of \mathbf{Z}_{ij} . Neither the intercept nor a separate vector for individual i is presented in the model because the same individuals have different options repeatedly (i.e., individual characteristics do not vary by options). Therefore, the choice probability is mainly defined by characteristics of the alternative neighborhoods in the model in Eq. (1). However, individual characteristics are included in the model as interaction terms because interactions between individual and neighborhood characteristics can vary. The statistical model that includes both individual and neighborhood characteristics can be expressed as follows:

$$P_{ij} = \exp(\mathbf{X}_i\beta_j + \mathbf{Z}_{ij}\alpha) / \sum_{k=1}^j \exp(\mathbf{X}_i\beta_k + \mathbf{Z}_{ik}\alpha). \quad (2)$$

Notice that the effect of individual characteristics \mathbf{X}_i vary with β_j , which is the log-odds of response j (i.e., selecting neighborhood j). Modeling the interactions between individual and neighborhood characteristics in this way, we can examine how neighborhood contexts matter differently for individuals with different characteristics. In other words, we can analyze how individual characteristics (such as nativity and SEI) are related to neighborhood characteristics (i.e., ethnic composition and SES of the place).

The structure of discrete choice models assumes that people could live anywhere and that they make choices from among all the available options. This possibility is, of course, not realistic; and as noted earlier, imposes a very heavy and perhaps unnecessary computational load. If we knew what people's actual choice set was, it would be preferable to model the process in two stages: first, to explain why people have different choice sets to begin with, and then to understand how they make specific choices within that range. An example could be based on distance from domicile to where a person works. Many people likely seek to minimize that distance, and if we knew where they worked (or where most people like them worked), we could profitably incorporate that information into the

model. In the absence of such data, we simplify to a single-stage process, not distinguishing between (1) places outside the choice set (because they are too far away or because the person couldn't afford it, wouldn't be allowed to live there, or knew nothing about it) and (2) places that were probably actually considered.

Findings

We present findings in four parts. First, we map the variations in street segments' occupation, nativity, family status, and race/ethnicity, and we report segregation measures based on these variables to assess the extent of residential differentiation at varying spatial scales. Second, we estimate discrete choice models at the scale of street segments, segment groups, and extended segment groups as another way to determine the spatial scale at which locational choices are made. Third, we assess the relative importance of class, nativity, family life cycle, and race/ethnicity in determining who lives where. Finally, we look more closely at individual coefficients and assess the size of their effects on estimated probability of living in a given neighborhood.

Segregation by Social Characteristic and Spatial Scale

Figures 2, 3, 4, and 5 provide maps of street segment characteristics in Newark in 1880. These maps offer a visualization of how people were spatially sorted. Figure 2 shows

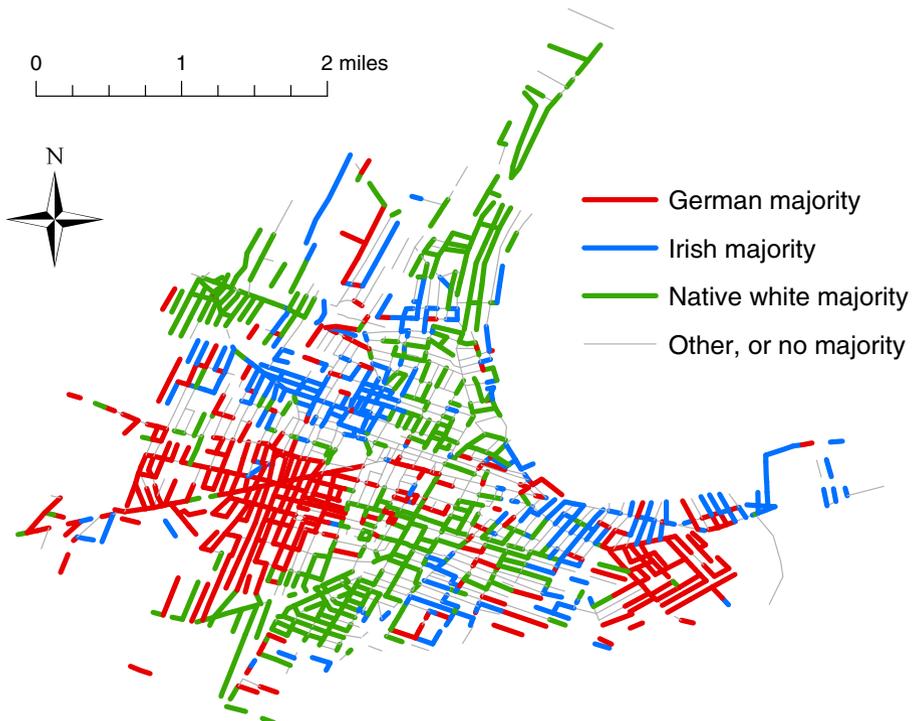


Fig. 2 Street segments by ethnicity

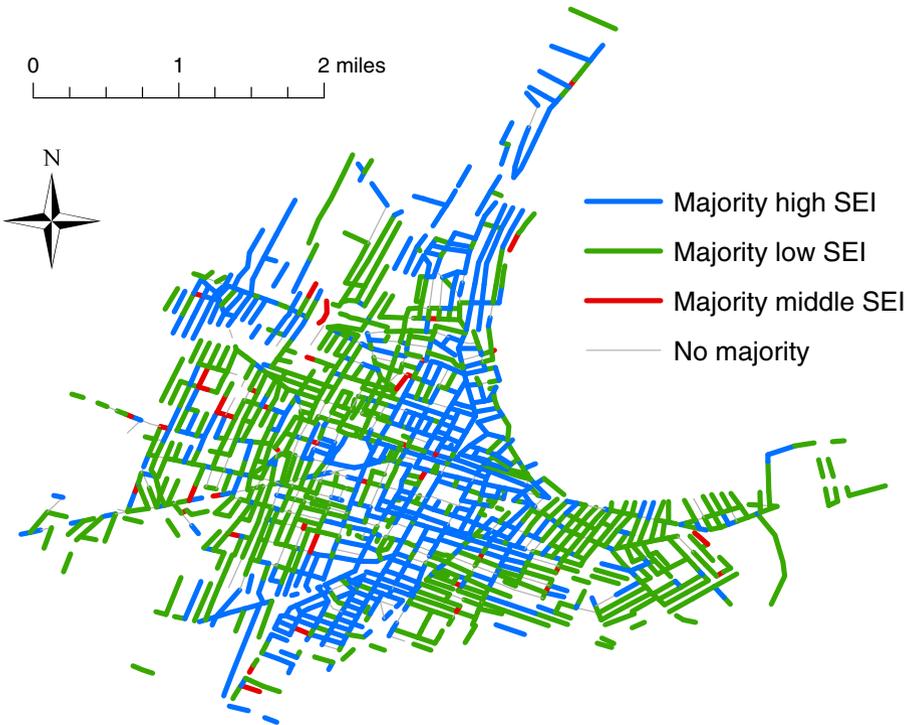


Fig. 3 Street segments by occupational standing

the majority ethnic group in street segments (native whites, Germans, and Irish, with a small number of segments that have no population or where another group or no group is a majority). Note the strong spatial clustering. Large adjacent zones are majority German or majority Irish in the west and southeast of the city. In between is a long corridor that is majority native white, stretching from the far north, through the central business district, and then southwest. Ethnic separation is apparent from this map, although it does not reveal to what extent street segments with a plurality of residents of one ethnicity may also include substantial minorities of another. Figure 3 illustrates segregation by occupational standing of the person with the highest SEI in each household on the street segment. SEI has been divided approximately into terciles. This map has some resemblance to the ethnic mosaic: areas with a majority of high-SEI households lie along the same north-south axis as native whites. Other areas are more spatially heterogeneous, with segments with no majority interspersed with those with a majority of low-SEI households. Figure 4 is also somewhat duplicative of the ethnic map because native white areas are quite likely to have a majority (shown here as 60 % or higher) of U.S.-born persons. However, within the predominantly German and Irish areas is a mixture of segments with varying levels of U.S.-born persons. Finally, Fig. 5 shows the distribution of households by family composition (the percentage of households that include a married couple living with their child of any age). There appears to be some spatial pattern in this map, but it is not nearly as clear as for the other characteristics, nor is it as closely linked to ethnicity.

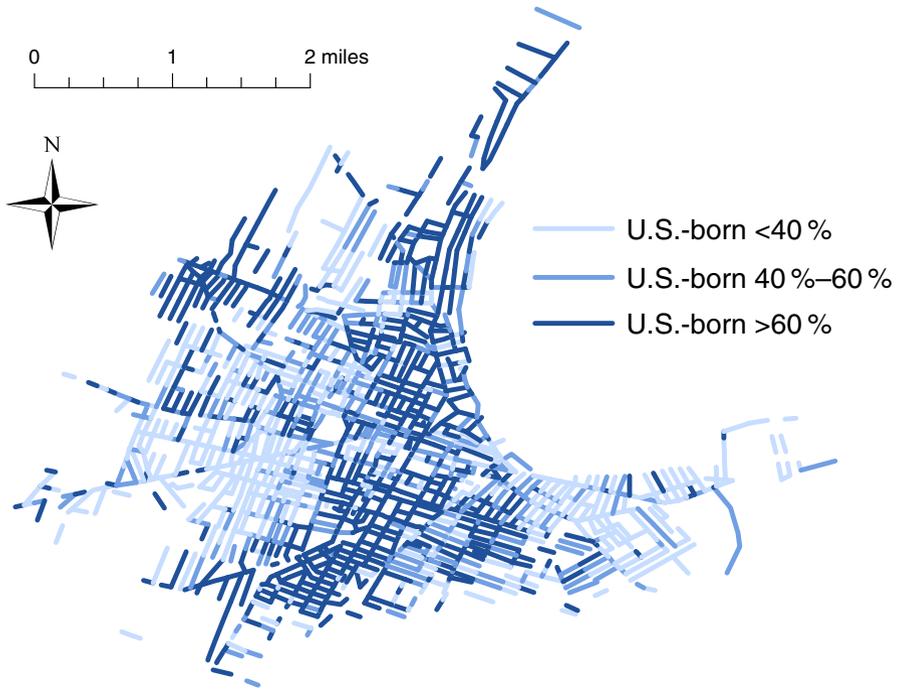


Fig. 4 Street segments by nativity

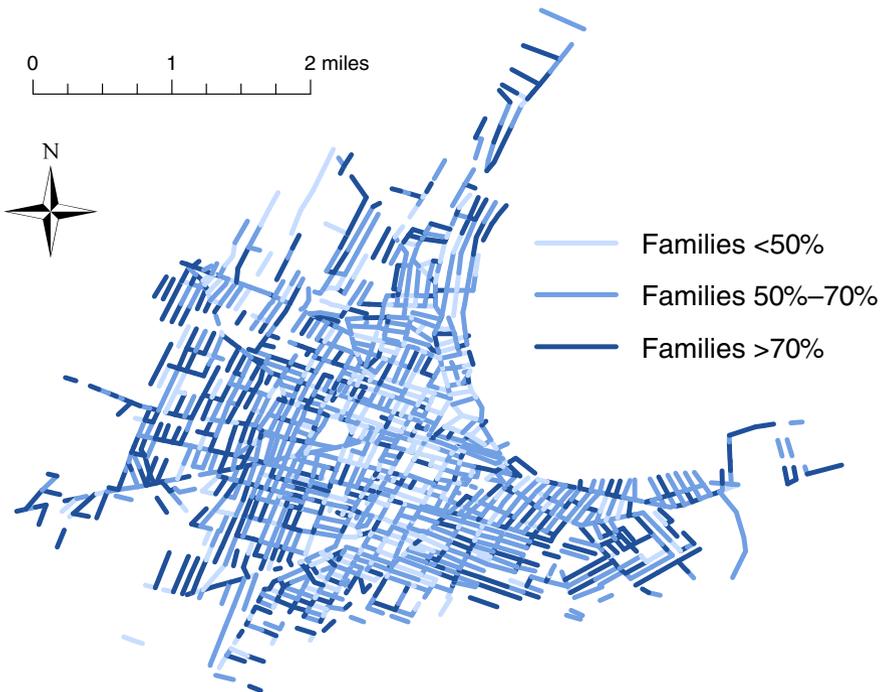


Fig. 5 Street segments by family status

Another way to assess how strongly people are sorted by a given characteristic and at what spatial scale is to calculate standard measures of segregation. We present two measures: the index of dissimilarity (D) and isolation (p^*). These measures are based on all 80,116 men and women age 18 and older who had valid occupational SEI scores. The index of dissimilarity is an overall measure of the extent to which persons in categories A and B are distributed in the same way across local areas. It achieves a value of 0 when the two distributions are the same, and has a maximum value of 1 when there is no overlap at all between the distributions. It requires that people be assigned to nonoverlapping local areas, and it has the advantage that it is not affected by the relative size of the two categories. Isolation, p^* , is a measure of homogeneity of local areas: the percentage of people in category A in the area where the average person in category A resides. This measure has the advantage that it can be defined for egocentric (and therefore overlapping) local areas.

Table 1 presents values of D across categories of people at the scale of street segments, segment groups, and extended segment groups. (These larger groupings are necessarily arbitrary; they were created to approximate, insofar as possible, the typical scale of segment groups and extended segment groups.) Note that people are much more segregated at a finer spatial scale, which is to be expected; at its limit, the “neighborhood” is the whole city, and at that spatial scale, there is by definition no segregation. What could not be known *a priori*, however, is how large a difference the scale makes and at what distances it matters most. Table 1 describes a fairly narrow range of spatial scales: extended segment groups extend no more than two blocks away from a focal street segment. We find that the largest decline in segregation occurs at the very beginning, between the street segment and the segment group, which includes people only one block away. For example, native white–German segregation in

Table 1 Levels of segregation (D) at various geographic scales

	Street Segment	Segment Group	Extended Segment Group
Race/Ethnicity			
Native white-black	.74	.53	.42
Native white-British	.37	.24	.20
Native white-German	.67	.59	.56
Native white-Irish	.57	.46	.41
Native white-others	.54	.42	.37
Occupational SEI (of valid cases)			
High vs. medium	.38	.31	.28
High vs. low	.37	.25	.21
Medium vs. low	.30	.20	.15
Family Status			
Single vs. married with children	.22	.17	.15
Nativity			
Native vs. foreign-born	.38	.32	.29

Notes: Sample is Newark men and women, age 18 and older. $n = 80,116$.

Newark in 1880 was .67 at the level of street segments, dropping to .59 for the slightly larger segment groups (an 8-point difference), and then to .56 for extended segment groups (only a 3-point decline).

Another new finding here is that levels of segregation also vary considerably depending on the social categories that are used. Racial/ethnic segregation tends to be highest. In a strong contrast to contemporary findings, the small black population of Newark was only slightly more segregated from native whites than were Germans at the level of street segments, and they were less segregated at larger scales. Segregation of native whites from British is quite low (.37 for street segments), and segregation of native whites from Irish is intermediate (.57 for street segments). Segregation by occupation and nativity is only moderate, and segregation by family status is low. This variation offers a clue as to what will turn out to be the more important bases of neighborhood formation.

Table 2 presents the measure of isolation. Isolation measures are strongly affected by the relative size of the category, so it is not advised to compare p^* values for different categories. Rather, the values of p^* are most interpretable in relation to the relative size of the category in the city (shown in column 1). The three largest ethnic categories (native whites, Irish, and Germans) were also the most isolated, living in street segments with approximately 50 % coethnic neighbors. In all these cases, isolation was much greater than expected simply on the basis of group size, which is due to the extent of their segregation. The same general phenomenon holds for categories of SEI, family status, and nativity; but on these dimensions, the level of isolation is much closer

Table 2 Levels of isolation (p^*) at different spatial scales

	Overall City Share	Street Segment	Segment Groups	Extended Segment Groups
Race/Ethnicity				
Native white	.28	.48	.42	.40
Black	.02	.18	.07	.05
British	.09	.17	.13	.12
Germans	.23	.55	.51	.49
Irish	.28	.45	.38	.35
Others	.09	.18	.13	.12
SEI (of valid cases)				
High	.30	.41	.37	.35
Middle	.47	.54	.52	.51
Low	.23	.30	.25	.23
Family Status				
Married with children	.42	.47	.44	.44
Other	.58	.61	.60	.60
Nativity				
Foreign-born	.46	.56	.53	.52
Native-born	.54	.63	.61	.60

Notes: Sample is Newark men and women, aged 18 or older. $n = 80,116$.

to people's share of the population. Again, this result is an indication that segregation on these dimensions was lower.

For race/ethnicity, again a large difference is evident across spatial scales, especially between segments and the other two scales. For example, the average black resident (in a city with only 2 % black population) lived in a street segment that was 18 % black, but this person's segment group and extended segment group were only 7 % and 5 % black, respectively. This finding means that there was a very sharp distance gradient in the clustering of black residents and that most black-white segregation would have been missed with data at the (now-standard) census-tract level. For other ethnic groups and categories of SEI, the gradient across scales is smaller but still evident. For instance, in a city where 30 % of persons had high SEI, high-SEI persons lived, on average, in segments that were 40 % high SEI, compared with segment groups that were 36 % and extended segment groups that were 35 %. However, there was almost no gradient by family status or nativity.

Discrete Choice Models at Different Spatial Scales

These aggregate analyses suggest tentative conclusions: (1) ethnicity was the primary basis of residential sorting, and (2) street segments were much more homogeneous than larger areas. We examine these points more closely through estimation of discrete choice models. As noted earlier, for this purpose, we use a random sample of men age 18 and older with valid occupational SEI, including younger men only if they were not living with their employed father. The full eligible population is 28,922. Table 3 provides the population counts for each racial/ethnic group and the total eligible population, along with means and standard deviations of the other individual-level measures. As noted earlier, the population was predominantly native white, German, and Irish. A very large majority of Germans, Irish, and British were foreign-born. Occupational SEI was markedly highest for native whites (37.6), and lowest for blacks (14.6) and Irish (22.5). Germans, British, and others were intermediate, at around 30. The majority of eligible persons were married with a child (of any age) living in the household. This share was highest for Germans (62 %) and lowest for blacks (42 %). Irish and British were near the average, while native whites and others were below the

Table 3 Characteristics of men eligible for sampling by race/ethnicity

	<i>N</i>	Foreign-born		Occupational SEI		Married With Children	
		Mean	SD	Mean	SD	Mean	SD
Native White	7,591	.00	.00	37.6	22.3	.53	.50
German	8,885	.85	.35	28.8	19.2	.62	.49
Irish	6,046	.79	.41	22.5	17.4	.56	.50
British	2,878	.72	.45	30.9	19.8	.57	.50
Black	665	.01	.10	14.6	12.5	.42	.49
Other	2,857	.59	.49	31.4	21.0	.49	.50
Total	28,922	.56	.50	29.9	20.6	.56	.50

Table 4 Model fit statistics for discrete choice analysis at three spatial scales

	-2 Log-Likelihood	Chi-Square	<i>df</i>	<i>p</i>
Street Segment	38,247	3,074	17	.001
Segment Group	38,682	2,357	17	.001
Extended Segment Group	38,906	1,956	17	.001

average. Not shown in the table, we also calculated average ages as a potential control variable given that some differences in SEI or family status could be affected by age differences across groups. The average age of sampled persons (all adults with a recorded occupation) was similar for all groups, ranging from 37.9 to 40.6 with a total mean age of 40.1.

The discrete choice models are based on a random sample of 2,894 persons living on 1,442 street segments. We begin by comparing the model fit at the three geographic scales. We include all contextual variables along with every interaction term linking a neighborhood characteristic with the corresponding individual-level characteristic. Table 4 summarizes the fit statistics. The table reveals a consistent increase in the log-likelihood statistic from smaller to larger neighborhood scales, indicating that the salience of a neighborhood characteristic in residential choice is gradually diminished with larger areas. The chi-square test values support the same conclusion. Because the three models have the same set of predictors ($df = 17$), their chi-square values are directly comparable as measures of goodness of fit. The largest chi-square value (3,074) is for the first model, suggesting that congruence of a person's own attributes with neighborhood characteristics measured at the level of street segments matters most for residential outcomes.

Strength of Predictors at the Street Segment Level

We turn now to the question of which combination of individual and neighborhood characteristics best predicts residential location. What is it about people and their potential neighbors that matters most? This analysis is presented only for street segments, although very similar results are found for segment groups and extended segment groups.

Measures of goodness of fit can be used in two ways to evaluate the relative strength of different categories of predictors. One is to estimate separate models in which only occupational SEI, only nativity, only family status, or only race/ethnicity are included. The other is to conduct a stepwise analysis in which each category of predictors is entered at a separate step. Table 5 summarizes results of both approaches.

Estimating separate models reveals that all these predictors are statistically significant. Race/ethnicity appears by far to be the most important, with a much higher chi-square (even considering the larger number of variables introduced) and much lower log-likelihood statistic. Family status appears to be the least important by the measures of model fit, with occupation and nativity in between.

In the stepwise analysis, we maximize the potential impact of occupation by entering it first into the model, and we minimize the effect of race/ethnicity by entering it last.

Table 5 Comparison of models predicting street segment choice by occupation, nativity, family, and race/ethnicity

	Overall				Change From Previous Step		
	-2 Log-Likelihood	Chi-Square	<i>df</i>	<i>p</i>	Chi-Square	<i>df</i>	<i>p</i>
Separate Models							
Occupation	40,442	408	2	<.001			
Nativity	40,183	651	2	<.001			
Family	40,676	173	2	<.001			
Race/ethnicity	38,980	2,500	11	<.001			
Stepwise Models							
Occupation	40,442	408	2	<.001	408	2	.001
Nativity	39,862	954	4	<.001	580	2	.001
Family	39,691	1,119	6	<.001	170	2	.001
Race/ethnicity	38,475	2,950	17	<.001	1,217	11	.001

Introducing occupational SEI in the first step results in a strong and statistically significant fit: people with higher SEI are much more likely to live on a street segment with higher mean SEI (not counting themselves). Chi-square for this model is 408. Adding nativity more than doubles the chi-square value. Immigrants are much more likely to live on street segments with higher proportions of foreign-born neighbors, and natives are more likely to live with other natives. Family status has a modest but statistically significant incremental effect. What stands out by far is the very large increase in chi-square and large reduction in log-likelihood associated with adding the race/ethnicity variables into the model.

Assessing the Magnitude of Effects

Table 6 presents the coefficients for each predictor and interaction term in the model. These are presented in pairs. The first term in each pair is the effect of the neighborhood characteristic by itself. The relevant term for our purpose is the interaction between an individual's attributes and the neighborhood characteristic. All the interaction terms are statistically significant, and all are in the direction predicted by homophily—that is, “like attracts like.” The Wald statistic measures the effect of the term on goodness of fit: how much would the fit be reduced if the term were omitted from the model? By this measure, homophily by occupational SEI and family status are more influential than nativity. The specific race/ethnicity terms vary in importance, and the order partly reflects the larger number of Germans and Irish in the sample.

Another way to assess variables' importance is by the size of *b* (or the exponent of *b*). In this respect, what stands out among race/ethnicity variables is the very large coefficient representing black segregation. However, all the race/ethnicity interaction terms are large and, as we found earlier, are very important in combination. To help assess the size of these effects, we show in Fig. 6 the variation in the predicted probability of living in a street segment according to the percentage of same-group

Table 6 Model coefficients for effects of occupation, nativity, family, and race/ethnicity on street segment choice

	<i>b</i>	SE	Wald	<i>df</i>	<i>p</i>	<i>exp(b)</i>
Occupation						
Mean SEI	-0.068	0.005	188.3	1	<.001	0.934
SEI × Mean SEI	0.176	0.011	247.7	1	<.001	1.192
Nativity						
Foreign %	-1.516	0.232	42.6	1	<.001	0.220
Foreign × Foreign %	2.012	0.218	85.3	1	<.001	7.480
Family						
Married with children %	-2.135	0.197	117.0	1	<.001	0.118
Married with children × Married with children %	3.051	0.247	153.1	1	<.001	21.146
Race/Ethnicity						
Native white %	0.531	0.374	2.0	1	.155	1.701
German %	0.633	0.389	2.7	1	.103	1.884
Irish %	0.890	0.387	5.3	1	.021	2.435
British %	1.268	0.418	9.2	1	.002	3.555
Other %	1.322	0.429	9.5	1	.002	3.752
Black % (omitted variable)						
Native white × Native white %	1.697	0.214	62.9	1	<.001	5.457
German × German %	2.921	0.158	343.5	1	<.001	18.558
Irish × Irish %	2.628	0.189	194.3	1	<.001	13.846
British × British %	3.188	0.403	62.4	1	<.001	24.237
Other × Other %	3.000	0.369	66.0	1	<.001	20.094
Black × Black %	5.717	0.597	91.7	1	<.001	303.895

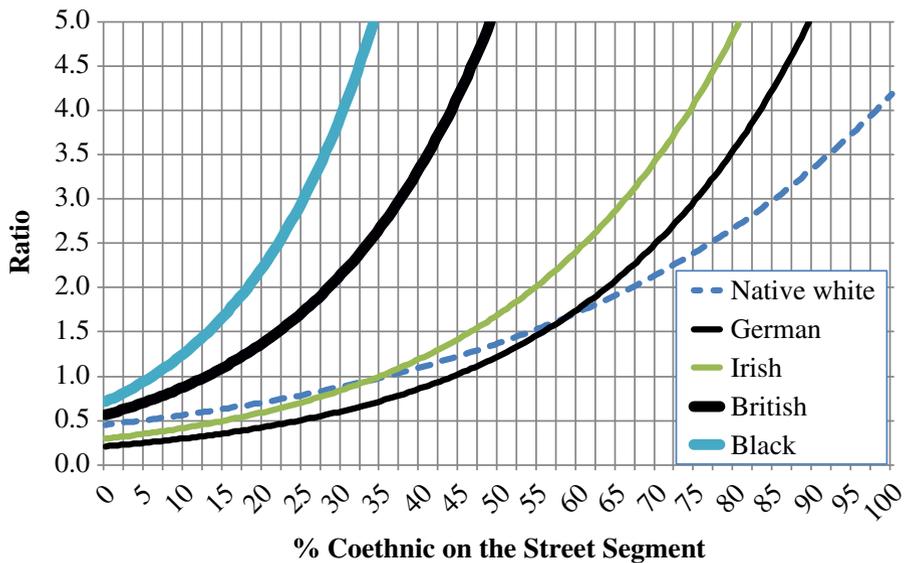


Fig. 6 Probability of living in a given street segment (ratio to a random placement) by share of coethnics on the segment

members residing in it. The y -axis is a probability ratio: the ratio of the predicted probability of living on a segment to the probability based on random assignment ($1/1,442 = .00069$). Values range from less than 1.0 (for low shares of coethnics) to 5.0 or higher (for high shares of coethnics). These curves are group-specific, and each one holds constant other variables in the model based on the average or modal values for respondents in that group.

Figure 6 shows that for blacks, the probability ratio rises rapidly with the share of blacks in the segment. Starting below 1.0, it reaches 3.0 at around 25 % black; by 35 % black, it is above 5.0. The British curve is also distinctive, reaching 1.0 at about 12.5 % British, then rising above 2.0 at 30 %. For the remaining groups, the rise is steady but more gradual up to about 50 % coethnic, which in fact represents approximately the average street segment on which these group members live (as shown in Table 2). After that point, the slope increases most for the Irish. The German curve is similar but somewhat more gradual. The smallest effect of coethnicity is for native whites.

Discussion and Conclusion

These findings should be interpreted with care. Although we frame the research question in terms of neighborhoods' spatial scale and their demographic basis, we have not actually identified neighborhoods in Newark in 1880. Rather, we have used the local area in which people live as a proxy for their neighborhood, assuming first that the neighborhood is the local street segment that they live on, then that the neighborhood also includes connected street segments, and then extending the boundaries out by another layer of connected segments. Thus, we have limited our study to egocentric neighborhoods, where a person is always near the geographic center. Newark may well have had some neighborhoods that were identifiable in other ways (such as homogeneous in class composition but with a diverse mix of other characteristics), or by some specific combination of class, ethnicity, nativity, and family status that we did not examine. These neighborhoods may have extended over many blocks, and some of our sample persons may have lived on the outer edge of the neighborhood instead of in its center. If we had known in advance what the "real neighborhoods" were, we would have approached our questions very differently. But as of yet, there is no consensus on how to identify neighborhoods. We have taken a step in that direction by providing some information about what seems to be the spatial scale of homophily (i.e., spatial sorting), which we take to be an important dimension of neighborhood. We have also made some progress toward determining the social characteristics from which neighborhoods are formed. However, real neighborhoods have other features that we did not study, such as a history and collective identity, local institutions, and organizations.

That said, our analysis shows that in Newark in this period, people lived near similar people at a very local scale, and that homophily declined significantly even a block away. This result is consistent with the intuition of urban ethnographers (such as Gerald Suttles) and with the perspective of more recent scholars (such as Rick Grannis) who have postulated that neighborhoods naturally build up from face-to-face interaction on a single street. Alternative findings were possible. For example, one could imagine a city where ethnic neighborhoods are well defined at a scale of dozens of city blocks. If non-ethnic residents were randomly distributed within the neighborhood, then any

given ethnic resident could well live on a somewhat diverse street segment just by chance. Yet, at a larger scale, such variations would fade. In that case, there could be as much information about coethnic neighboring from data at a multiblock scale as at the scale of a street segment. That is not what we found.

One might suspect that our findings on spatial scale could have resulted from the fact that sampled persons themselves—because they live on the street segment—affect its composition. In that case, the stronger associations between personal and neighborhood characteristics would be built in by the methodology. However, because we have full count data for Newark, we have been able to measure neighborhood composition without including the sampled person. Characteristics of the sampled person cannot be thought of as “causing” the composition of the neighborhood in this analysis.

Is this result connected with the period that we study? Possibly so. In 1880, most working-class people probably walked to work. The available horse-drawn trams were slow, expensive, and designed especially to serve more affluent people who lived in newly developing suburbs. People likely did their routine shopping very close to home, creating relatively small micro-environments of daily living. Possibly as mass public transportation and use of private cars became important, people tended to move across greater distances and their neighborhoods both expanded in geographic scale and became more distant from one another. As geocodable census microdata from a more recent period become available, or as large databases emerge that track people’s movements through smart phones and other devices, we will gain a capacity to study this question. We believe that discrete choice models such as the one used here will be informative about historical changes in the scale of neighborhood formation, and much can be learned by applying them to both historical and contemporary data.

It may be argued that for the purpose of studying spatial inequality, it doesn’t matter that sorting occurs mainly at the scale of street segments rather than larger units. As noted earlier, Suttles (1972) believed that larger areas constituted what he called “defended neighborhoods” in which one could expect to find collective mobilization. It is at these larger scales that school segregation and other public infrastructure are often determined, and for some purposes, the “real” boundaries of neighborhoods are predetermined by service boundaries or electoral boundaries or other similar sharp dividing lines.

Yet, spatial inequality is not the only motivation for studying segregation patterns. In fact people’s choices about where to live are strongly constrained by conditions at the most local scale, then studying choices or outcomes at that scale is also sociologically important. Choices at that scale—the locale within which people are most likely to have face-to-face interactions and form neighborhood-based social ties—may reveal most clearly the nature of social boundaries. Possibly even if formal services and political representation are organized at larger scales, other resources (such as social support networks, friendship patterns, or networks of information about jobs) may be organized at the same scale as choices about where to live. Newark’s 1880 neighborhoods were formed as individuals located themselves among similar neighbors on a single street segment. We don’t know whether the same is true today, but our point is that how neighborhoods are formed matters to our understanding of residential patterns.

Our other major finding is that several different social characteristics have a role in people’s residential placement. This, too, is a question for comparative studies, and discrete choice models can again be a valuable tool. Unlike studies of segregation that

evaluate spatial differentiation along one social characteristic at a time, the discrete choice model allows us to deal with them simultaneously. Both the data demands (large samples of geographically referenced microdata) and the computing load are high, but we can anticipate that both will become less limiting in the near future. Ethnicity stands out as a determinant in the Newark case (as does race, despite the very small black presence in the city). Class (as indicated by occupational standing), nativity, and family status also are significant.

Again we can ask whether the predominance of ethnicity is a function of the period or the groups being studied. Urban historians Warner (1962) and Zunz (1982) explicitly argued that high levels of ethnic segregation in 1880 were subsequently replaced by growing segregation by social class. Quillian (2015:258), applying discrete choice models to contemporary data, reported that, “When neighborhood race and income are considered together, most of what appears to be race sorting into neighborhoods of different income levels is actually race sorting by racial composition.” At the height of early suburban expansion after World War II, new suburbs were heavily dominated by single-family homes oriented toward the needs of families with children. Hence, family status may have become more significant at some point. Our findings for one city in 1880 raise questions and illustrate an approach to studying how neighborhood formation varies across urban areas and over time.

Acknowledgments Earlier versions of this article were presented at the Eastern Sociological Societies, April 2010, and RC28 meeting at Chinese University of Hong Kong, May, 2012. This research was supported by research grants from National Science Foundation (0647584) and National Institutes of Health (1R01HD049493-01A2). The Population Studies and Training Center at Brown University (R24 HD041020) provided general support. The authors have full responsibility for the findings and interpretations reported here.

References

- Alba, R., & Logan, J. (1992). Analyzing locational attainments: Constructing individual-level regression models using aggregate data. *Sociological Methods & Research*, 20, 367–397.
- Alba, R., Logan, J., & Stults, B. (2000). How segregated are middle-class African Americans? *Social Problems*, 47, 543–558.
- Anderson, E. (1992). *Streetwise: Race, class, and changes in an urban community*. Chicago, IL: University of Chicago Press.
- Berry, B. J. L., & Rees, P. (1969). The factorial ecology of Calcutta. *American Journal of Sociology*, 74, 445–491.
- Bider, F., & Reimers, D. (1995). *All the nations under heaven: An ethnic and racial history of New York City*. New York, NY: Columbia University Press.
- Bruch, E. (2014). How population structure shapes neighborhood segregation. *American Journal of Sociology*, 119, 1221–1278.
- Bruch, E., & Mare, R. (2006). Neighborhood choice and neighborhood change. *American Journal of Sociology*, 112, 667–709.
- Bruch, E., & Mare, R. (2012). Methodological issues in the analysis of residential preferences and residential mobility. *Sociological Methodology*, 42, 103–154.
- Burgess, E. W. (1925). The growth of the city: Introduction to a research project. In R. E. Park, E. W. Burgess, & R. McKenzie (Eds.), *The city* (pp. 1–46). Chicago, IL: University of Chicago Press.
- Chaix, B. (2009). Geographic life environments and coronary heart disease: A literature review, theoretical contributions, methodological updates, and a research agenda. *Annual Review of Public Health*, 30, 81–105.

- Charles, C. (2003). The dynamics of racial residential segregation. *Annual Review of Sociology*, 29, 167–207.
- Coulton, C. J., Korbin, J., Chan, T., & Su, M. (2001). Mapping residents' perceptions of neighborhood boundaries: A methodological note. *American Journal of Community Psychology*, 29, 371–383.
- Denton, N., & Massey, D. (1988). Residential segregation of blacks, Hispanics, and Asians by socioeconomic status and generation. *Social Science Quarterly*, 69, 797–817.
- Dietz, R. D. (2002). The estimation of neighborhood effects in the social sciences: An interdisciplinary approach. *Social Science Research*, 31, 539–575.
- Drummond, J. O. (1979). *Transportation and the shaping of the physical environment in an urban place: Newark, 1820–1900* (Unpublished doctoral thesis). New York University, New York, NY.
- Duncombe, W., Robbins, M., & Wolf, D. (2001). Retire to where? A discrete choice model of residential location. *International Journal of Population Geography*, 7, 281–293.
- Erickson, A. J. (1995). Moving day. In K. T. Jackson (Ed.), *The encyclopedia of New York City* (pp. 778–779). New Haven, CT: Yale University Press.
- Galster, G. (2001). On the nature of neighbourhood. *Urban Studies*, 38, 2111–2124.
- Grannis, R. (2009). *From the ground up: How the layered stages of neighbor networks translate geography into neighborhood effects*. Princeton, NJ: Princeton University Press.
- Hershberg, T. (Ed.). (1981). *Philadelphia: Work, space, family and group experience in the nineteenth century*. New York, NY: Oxford University Press.
- Hoffman, S., & Duncan, G. (1988). Multinomial and conditional logit discrete-choice models in demography. *Demography*, 25, 415–427.
- Hunter, A. (1974). *Symbolic communities*. Chicago, IL: University of Chicago Press.
- Iceland, J., & Wilkes, R. (2006). Does socioeconomic status matter? Race, class, and residential segregation. *Social Problems*, 53, 248–273.
- Jargowsky, P. (1996). Take the money and run: Economic segregation in U.S. metropolitan areas. *American Sociological Review*, 61, 984–998.
- Kaplan, D. H., & Holloway, S. R. (2001). Scaling ethnic segregation: Causal processes and contingent outcomes in Chinese residential patterns. *GeoJournal*, 53, 59–70.
- Kusenbach, M. (2008). A hierarchy of urban communities: Observations on the nested character of place. *City & Community*, 7, 225–249.
- Lee, B. A., Reardon, S. F., Firebaugh, G., Farrell, C. R., Matthews, S. A., & O'Sullivan, D. (2008). Beyond the census tract: Patterns and determinants of racial segregation at multiple geographic scales. *American Sociological Review*, 73, 766–791.
- Lieberman, S. (1963). *Ethnic patterns in American cities*. New York, NY: Free Press.
- Logan, J. R., Alba, R., & Zhang, W. (2002). Immigrant enclaves and ethnic communities in New York and Los Angeles. *American Sociological Review*, 67, 299–322.
- Logan, J. R., Jindrich, J., Shin, H., & Zhang, W. (2011). Mapping America in 1880: The Urban Transition Historical GIS Project. *Historical Methods*, 44, 49–60.
- Logan, J. R., Stults, B., & Farley, R. (2004). Segregation of minorities in the metropolis: Two decades of change. *Demography*, 41, 1–22.
- Mare, R., & Bruch, E. (2003). *Spatial inequality, neighborhood mobility, and residential segregation* (California Center for Population Research Working Paper, No. 003-03). Los Angeles: University of California, Los Angeles.
- Massey, D. (1985). Ethnic residential segregation: A theoretical synthesis and empirical review. *Sociology and Social Research*, 69, 315–350.
- Massey, D., & Denton, N. (1993). *American apartheid*. Cambridge, MA: Harvard University Press.
- McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in economics* (pp. 105–135). New York, NY: Wiley.
- McFadden, D. (1978). Modeling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickers, & J. W. Weibull (Eds.), *Spatial interaction theory and planning models* (pp. 75–96). Amsterdam, The Netherlands: North-Holland.
- McKenzie, R. D. (1968). The neighborhood. In R. D. McKenzie (Ed.), *On human ecology* (pp. 73–83). Chicago, IL: University of Chicago Press. (Original work published 1921)
- O'Sullivan, D. (2005, November). *Impacts of neighbourhood scale on the dynamics of segregation*. Paper presented at the 17th Annual Colloquium of the Spatial Information Research Centre, University of Otago, Dunedin, New Zealand. Retrieved from http://www.business.otago.ac.nz/sirc/conferences/2005/06_osullivan.pdf
- Popper, S. H. (1952). *Newark, N. J. 1890–1910: Chapters in the evolution of an American metropolis* (Unpublished doctoral thesis). New York University, New York, NY. (Original work published 1921)

- Quillian, L. (2015). A comparison of traditional and discrete-choice approaches to the analysis of residential mobility and locational attainment. *Annals of the American Academy of Political and Social Science*, 660, 240–260.
- Reardon, S., Matthews, S. A., O'Sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The geographic scale of metropolitan racial segregation. *Demography*, 45, 489–514.
- Sampson, R. J., Raudenbush, S., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277, 918–924.
- Schelling, T. (1978). *Micromotives and macrobehavior*. New York, NY: Norton.
- Scott, A. J. (1990). *Metropolis: From the division of labor to urban form*. Berkeley: University of California Press.
- Shevky, E., & Bell, W. (1961). Social area analysis. In G. A. Theodorsen (Ed.), *Studies in human ecology* (pp. 226–335). New York, NY: Harper & Row.
- Sobek, M. (1996). Work, status, and income: Men in the American occupational structure since the late nineteenth century. *Social Science History*, 20, 169–207.
- Spielman, S. E., & Logan, J. R. (2013). Using high-resolution population data to identify neighborhoods and establish their boundaries. *Annals of the Association of American Geographers*, 103, 67–84.
- Suttles, G. D. (1972). *The social construction of communities*. Chicago, IL: University of Chicago Press.
- Van Arsdol, M. D., Jr., Camilleri, S. F., & Schmid, C. F. (1958). The generality of urban social area indexes. *American Sociological Review*, 23, 277–284.
- Warner, S. B. (1962). *Streetcar suburbs: The process of growth in Boston 1870–1900*. Cambridge, MA: Harvard University Press.
- Wepman, D. (2002). *Immigration: From the founding of Virginia to the closing of Ellis Island*. New York, NY: Fact on File.
- White, M. (1987). *American neighborhoods and residential differentiation*. New York, NY: Russell Sage Foundation.
- Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Cambridge, MA: Harvard University Press.
- Zunz, O. (1982). *The changing face of inequality: Urbanization, industrial development, and immigrants in Detroit, 1880–1920*. Chicago, IL: University of Chicago Press.