Journal Title
XX(X):1–24

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

The Dynamics of Urban Neighborhoods: A Survey of Approaches for Modeling Socio-Spatial Structure

Abstract

For close to a century, researchers from across the disciplines of Urban Studies have developed empirical models for understanding the spatial extent and social composition of urban neighborhoods–and how these dimensions change over time. Unfortunately, however, these techniques have often been developed within disciplinary silos and without broad exposure to other potentially interested constituencies. In this paper, we traverse the literatures of social science, computer science, and statistics to examine a variety of modeling techniques for understanding neighborhood dynamics. We begin our review by examining early concepts of spatial structure first outlined in the Chicago School and discuss how the notions of social ecology and quantitative neighborhood analysis permeated the urban studies for several decades to come. Our survey continues by reviewing contemporary statistical approaches for identifying urban neighborhoods, culminating with the state of the art in subfields known as 'geodemographics' and 'regionalization'. Following this review, we offer insight into the field's persistent conceptual issues, identify areas ripe for additional research, and highlight newly-developed computational methods that can inform more just and socially equitable public policy, community development, and accountable governance.

Keywords

Neighborhoods, Geodemographics, Regionalization, Factorial Ecology, Social Area Analysis, Ecometrics

1 Introduction

The analysis of neighborhoods - including their demographic makeup, physical morphology, environmental conditions, and the evolution of these features over time - is a central focus of research in both the social and physical sciences. Indeed, for nearly a century, neighborhoods have been among the most prominent primitive units of analysis in the social sciences, owing largely to important and durable contributions to urban theory made by the early Chicago School (Park et al. 1925; Merriman 2015). Park, Burgess, and McKenzie understood the centrality of space in social outcomes and set out to develop theories and models to explain the social fragmentation they observed in Chicago. Ever since, social scientists have championed their cause tirelessly, attempting to disentangle the reciprocal nexus linking cities, people, and neighborhoods. Yet despite their ubiquity in the scholarly literature, neighborhoods remain enigmatic. There is no precise definition of "neighborhood" in either spatial extent or social composition, and the continued use of the term belies its fundamental complexity, both in concept and empirical operationalization. Put simply, neighborhood research is difficult for many reasons but necessary for many more. Neighborhoods are fundamental elements of social life, and their spatial configurations have deep implications for the human experience.

In the contemporary era, there is a vast and renewed interest in empirical neighborhood analysis. Substantively, this push is driven by two social trends. The first is a growing recognition of the importance and pervasiveness of "neighborhood effects" in shaping social inequality and helping to produce a wide variety of stratified outcomes in areas like health (Diez Roux 2001; Diez Roux and Mair 2010), educational attainment (Garner and Raudenbush 1991; Burdick-Will et al. 2010), cognitive development (Sampson et al. 2008; Sharkey and Elwert 2011), employment (Mendenhall et al. 2006; Galster 2017), and economic mobility (Chetty et al. 2014, 2015), among a wide variety of others (Sampson et al. 2002; Sampson 2012a; Galster 2012; Sharkey and Faber 2014; Knaap 2017; Galster and Sharkey 2017). The second is the rise of "data science," and computational research methods, particularly the growing subfield of geographic or spatial data science, and the increasing adoption of advanced quantitative techniques for studying urban areas. Indeed, Wolf (2018) has even gone so far as to say that the identification of neighborhood clusters is the "quintessential geographic data science problem."

In the following review, we discuss the convergent developments across distinct strands of literature. Among these bodies of scholarship, two run parallel to one another: the first is concerned with developing typologies of places in which the residents who inhabit those places are internally similar according to their various attributes (a practice known as geodemographics). Because they are derived from many underlying variables, the neighborhood typologies are conceived as a type of latent characteristic that captures socio-spatial context. These classifications are then taken as true measures of this socio-spatial context and used as data in secondary models of phenomena of interest.

The other neighborhood analysis literature leverages similar statistical techniques, albeit with a different intent. Whereas the first literature is inspired by the Chicago school's understanding of urban spatial demarcations through sociodemographic segregation, the second literature is inspired by geographic data mining and computer science. Instead of estimating and mapping these purely-social distinctions, *urban regional science* detects socio-spatial "regions," the spatially-coherent places latent within sociodemographic data. Thus, urban regional science sees geography as the *embodiment* of sociodemographic segregation, and thus is an object of study itself, rather than a medium of how sociodemographic segregation could be visualized or *expressed*. This embodied-expressed distinction is fundamental to the difference between the two approaches.

To facilitate a better understanding of these two perspectives in the vast literature on neighborhood identification, classification, and transition, we review the last 100 years of research focused on measuring and modeling urban socio-spatial structure. In so doing, we discuss the empirical foundations and theoretical underpinnings underlying each approach; we highlight the prototypical examples as well as the innovations and pitfalls of each method, and finally, we posit new avenues for conceptual and computational advances.

2 Expression & Embodiment: Two Perspectives on Urban Space

First, we discuss the Chicago School and its durable contribution to the understanding of urban space. The Chicago School is an important anchor not because its scholars were the first to study spatial structure, nor because their theories were infallible, but because the School's founding marks the symbolic start of the spatial turn in the social sciences; scholars even today credit the Chicago School as the inspiration for neighborhood analysis writ large. After setting forth this perspective, we examine the development of urban regional science.

The Chicago School: Expressing Sociodemographic Space

The Chicago School generally traces its genesis to "The City," the seminal volume compiled by Park et al. (1925) that laid out the central tenets of a new brand of urban sociology. This new scholarship was a response to the changing human condition evidenced shortly after the turn of the 20th century, and the waves of immigration, urbanization, and increasing heterogeneity beginning to characterize the United States and other developed countries in the West (Wirth 1938). Led by Park, Burgess and McKenzie, the early Chicago School was characterized by "a focus on the city, concern for the effects of urbanization in producing social disorganization, an interest in the breakdown of the extended family and the neighborhood as a social unit, a consideration of the conflict between transplanted old world cultures and the emergent culture of new American cities, and a commitment to empirical methods" (Bell and Greer 1962). This shift in focus toward the urban social experience was a noteworthy turn in sociology, not only because "it is one of the most well-known turning points in…social science research," but also because it signaled the growing appreciation for the effects of spatial structure, and "brought neighborhood-centered research to the fore of the discipline during the early twentieth century" (Sampson 2012a).

Indeed, the focus on spatial structure was so significant that the "Chicago [school] felt that no social fact makes any sense abstracted from its context in social (and often geographic) space and social time. Social facts are *located*" (emphasis original) (Abbott 1997, p.1152).

To make sense of this relationship between people and places, the Chicago school scholars drew upon natural ecology, borrowing the concepts of invasion and succession to explain patterns of residential mobility among different demographic and socioeconomic groups. Examining Chicago's segregated landscape, Park and his colleagues set out to describe the geographic pattern they observed in the city, arguing that its residents were in a constant state of competition for space. As residents climbed the social ladder, they would trade economic capital for spatial capital, and reap the positive benefits of new pro-social communities (McKenzie 1924; Park et al. 1925; Wirth 1938; Fischer 1972). The process is summarized elegantly by Logan (1978, p. 407): “Residential segregation creates a status hierarchy of neighborhoods defined simply by the characteristics of their residents, at the same time as common class or status becomes a symbol through which people identify their physical area as a community. The status hierarchy of places is reinforced by people’s individual decisions to translate upward social mobility into change of place of residence”. In the parlance of Chicago School adherents, urban dwellers engage in processes of spatial assimilation, and the social structure of the city leads to a predictable spatial pattern organized in concentric rings extending from the center of the city (Burgess 2008). Park and Burgess based their model on the early economic work of Von Thunen’s “Isolated State”, which argued that transportation costs were the central organizing factor in urban economics, and that willingness to pay for access to the city center would lead to a predictable partitioning among agricultural, residential, commercial, and industrial land uses - eventually leading to the classical “bid-rent” model of urban economics (Alonso 1964; Mills 1967; Sinclair 1967; Muth 1969).

The Chicago School is thus an important frame for understanding contemporary approaches for modeling spatial structure, not only because its main conceptual contribution, the Burgess Concentric Zones (BCZ) model, is the standard-bearer, but because the same kinds of demographic shifts are now taking place that once caused urban scholars to engage deeply with spatial structure: globalization is shaking up both migration and urban economic structure. Attention to growing socioeconomic inequality has once again brought neighborhood effects to the fore of urban studies, and gentrification together with the “back to the city” movement are reshaping geographies of segregation and inclusion. Thus, the time is ripe to re-examine a series of questions guided by the Chicago School tradition.

The first question surrounds the axes of neighborhood differentiation: what are the relevant dimensions that distinguish one neighborhood from another? Are they physical (e.g. walls, train tracks, highways)? Are they social (e.g. defacto segregation)? How are these boundaries maintained through social and political processes? How are the boundaries changing? How does the maintenance of social and spatial boundaries lead to the development of “community” and the socialization of certain types of behaviors?

Second, how can (and should) neighborhoods be operationalized and urban spatial structure analyzed: What are the commonly used techniques for modeling spatial structure? Are the methods inductive or deductive? What are the challenges and pitfalls associated with modeling spatial structure? How can models of spatial structure be used to develop better, more equitable urban policies?

Modelling Social Structure

Following the introduction of BCZ, empirical models of spatial structure went largely undeveloped for a significant period. In the 1950s, however, social shifts yet again brought emphasis on the significance of place. During the postwar period of suburbanization, white flight, and social unrest, there was a significant push to understand the nature of “community;” the Chicago School focused strongly on how urban spatial segmentation leads to social behaviors like territorialism and consciously created “ideological communities” (Hawley 1950; Suttles 1972; Hunter 1975). To understand how these communities were created, scholars turned to newly developed statistical methods to identify the essential elements of “urbanism” that structure modern life and the field of social area analysis was born. This turn is often viewed as the beginning of an age of urban empiricism, but it is important to emphasize that a critical component of Chicago School analysis is empirical work that is grounded firmly in social theory (Sampson et al. 2002). As researchers adopted new statistical techniques, therefore, they attempted to operationalize “the complicated phenomena of urbanism,” described by Wirth (1938, p. 19) as “a system of social organization involving a characteristic social structure, a series of social institutions, and a typical pattern of social relations”.

Seeking to formalize and operationalize the ideas of human ecology, neighborhoods, and social areas, researchers set out to define neighborhoods as a set of spatially structured social interactions. A neighborhood or “natural area” could then be identified as meeting the following criteria:

(a) a geographic area physically distinguishable from other adjacent areas;
(b) a population with unique social, demographic, or ethnic composition;
(c) a social system with rules, norms, and regularly recurring patterns of social interaction that function as mechanisms of social control; and

(d) aggregate emergent behaviors or ways of life that distinguish the area from others around it" (Schwirian 1983, p. 84)

These ideas set the stage for an entire generation of researchers focused on discovering the latent spatial structure in social relations. What is particularly important about the framework Schwirian articulates is encapsulated in the last two bullets. The Chicago School and its devotees maintained a focus on human behavior, cultural norms, and assimilation, which they viewed as having a reflexive relationship with residential arrangements. Early approaches were therefore driven by a focus on identifying, isolating, and quantifying the social processes that led to territorialism, "defended communities", and segregation, among other emergent behaviors (Suttles 1972).

One of the earliest innovations in neighborhood empirical work was the notion that social processes, which are difficult to observe, could be treated as latent variables and modeled using easily obtained Census data, similar to the ways that psychologists were beginning to model unobservable personality traits in individuals. The first studies deploying this technique were known as Social Area Analyses (SAA), and were developed to help understand the shifting patterns of segregation and urbanization that began in the 1950s. Although social area analysis has been long studied and its lineage is well-known, it is important to remember that its early emphasis on natural science and social psychology generated a empirical search for the fundamental axes of community differentiation–the laws of social physics that described how segregation and city living restructured the life course.

Social Area Analysis Social area analysis was first devised by Shevky and Williams (1949) and uses factor analysis to isolate and measure what the authors conceived as three essential dimensions of urban spatial structure:

1. *urbanization* - measured by manifest variables fertility, women in the labor force, and single-family dwelling units
2. *social rank* - measured by manifest variables occupation, educational attainment, and rent
3. *segregation* - measured by an "index of isolation for selected ethnic and foreign-born groups" (Bell 1953)

Together, Shevky and Bell postulated, these three constructs accounted for the majority of the differences between population groups living in the city. The Shevky-Bell hypothesis, as it is now known, holds that urbanization, segregation, and social rank are the defining forces that structure urban life and influence a variety of behaviors like household formation and participation in formal organizations (Bell 1953; Spielman and Thill 2008). Implicit in SAA is that these factors have theoretical connections to behavior. Urbanization, for example may lead to declining birth rates as women stop bearing children and join the workforce, and segregation may lead to predictable residential patterns as immigrants and ethnic minorities form enclaves for mutual benefit*. Following their identification, the latent factors are used as input to cluster analysis, used to group neighborhoods into similar types. Describing Shevky's original conceptual framework for SAA, Herbert (1967, p. 42) articulates a case that modern urban industrialism is characterized by unavoidable "changes in the distribution of skills, changes in the organisation of productive activity, and changes in the composition of population. Associated with these three main trends are the expressions of social differentiation which become more marked over time. Thus, Shevky & Williams's SAA specifies three goals: first, SAA specifies a quantitative framework for capturing these three essential dimensions of social transformation; second, SAA finds groups of spatial units (neighborhoods, in theory) that are similar along each of the three dimensions. By grouping the neighborhoods into categories, Shevky hoped to capture nonlinear dynamics that might result from the interaction of the three components. Third, SAA uses the resulting "social area" categories as lenses and explanatory variables for other urban inquiries (Brindley and Raine 1979).

Shevky and Williams's initial work focused on Los Angeles, and soon after it was published, Bell (1953) reimplemented SAA in San Francisco, using his results first to examine the generalizability of the original L.A. study, (Bell 1955) and later to study spatially stratified participation in organizations, and informal social relations in different neighborhood types (Bell and Force 1956; Bell and Boat 1957). A number of replications were also performed to test the stability of the Shevky-Bell hypothesis, and whether the same general structure appeared in other American cities, which it often did (Schmid 1950; Greer 1956; Schmid et al. 1958; Arsdol et al. 1958b,a).

Despites some converging results from different cities, the replication studies often were contentious. Some objected to the use of SAA, arguing that it lacked foundations in social theory and, apart from interesting patterns, provided

*Later, SAA and human ecology more broadly would be criticized for misunderstanding the causal nexus of these patterns (e.g. that segregation is not entirely due to choice), but here the important point is that variable selection in SAA was driven by attention to social processes, rather than by simply interesting or available variables little insight into the causes and consequences of urban ills (Hawley and Duncan 1957; Van Arsdol, et al. 1961; Van Arsdol et al. 1962). In a particularly poignant critique, Hawley and Duncan (1957) question whether urbanization, segregation, and social rank are the defining characteristics of cities, and whether the social areas analysis of these variables provides valuable insight into urban life. Put bluntly, they argue that the 'social area' lacks scientific rigor because it "has provided no theory that explains why areas tend to be homogeneous or otherwise, or that predicts the degree of homogeneity to be observed" [p.339].

Apart from methodological issues, some replications called into question whether the three factor structure was a sufficiently robust model to describe American cities. While there was general support for the SAA model, some results were mixed, particularly with respect to the strength and orthogonality of the three factors, leading Anderson and Bean (1961)

to question whether the same factor solution would emerge using alternative variables and whether the factors would remain static in number and interpretation. Anderson and Egeland (1961) probe the question in more depth, finding support for Burgess's concentric zones theory in terms of urbanization but not with respect to social rank, and Udry (1964) argues that the factor solution is sensitive to the size of the spatial units. In the ongoing debate, even Bell and Greer (1962) conceded that while "there is clearly emergent and presumptive evidence of verified theoretical structure in the Shevky schema of urban analysis, additional specifications, elaboration, and formulation are necessary," and as more human ecologists heeded his call, exploratory investigations of the factor structure of urban areas blossomed into their own subfield called "factorial ecology."

Factorial Ecology Through the 1970s a staggering number of Factorial Ecology (FE) studies appeared in the literature. Unlike its predecessor SAA, which tried to derive three theoretically meaningful constructs using factor analysis, then performed a cluster analysis on those axes to understand urban segmentation, FE is typically more open-ended. Instead, FE is an inductive approach that leverages exploratory factor analysis and eschews clustering. Similar to Anderson and Bean (1961), factorial ecology researchers are interested in how the urban social structure might be modeled if a more diverse set of variables were factored. Following, in FE many social variables are provided to a factor model and components emerge from the data. By examining which variables load strongly on which factors, researchers can intuit the conceptual interpretation of each component. Conceptually, factorial ecology borrows from psychological personality research and psychometrics, modifying "factors of the mind" into factors of the neighborhood (Palm and Caruso 1972). Berry (1971) and Rees (1971) develop "factor

models" that borrow from psychological personality research and factor labeling in factorial ecology. Factor models are sensitive to the choice of rotation (oblique or orthogonal) and the estimation procedure used (Hunter 1972). Accordingly, many of the results that have emerged from studies of factorial ecology should be treated with skepticism until it can be shown that they are robust to the choice of factor model used (Salins 1971; Taylor and Parkes 1975; Newton and Johnston 1976; Perle 1979). Others have criticized factorial ecology for lacking theory, arguing that it amounts to quantitative fishing, since any derived factor structure can be explained ex-post-facto.

Despite these criticisms, FE studies have been undertaken all over the world and its diverse applications were even featured in a special issue of Economic Geography (Berry 1971). Scholars have canvassed a wide variety of continents, cultures, and class-systems, including Ireland (Parker 1975) Sweden (Janson 1971), the Middle East (Landay 1971), India (Rees 1969; Berry 1971), Canada (Murdie 1969; Bourne and Barber 1971), and Brazil (Morris and Pyle 1971), in addition to Los Angeles, Chicago (Hunter 1971), a wide variety of other American cities (Palm and Caruso 1972), and more. Despite this broad applicability–or perhaps because of it–Landay (1971) raises the issue of generalizability and cultural sensitivity when applying factorial ecology in different contexts. Of particular interest is the method by which variables are chosen. If the focus is on a "contextual mode of inquiry," how much hyperlocal context needs to be embedded in the data, and how "standardized" might be the results? Perfect contextual data is infinitely nuanced by definition, and any attempt to distill a set of "standardized" results is, therefore, off the mark. Thus "if the goal is to make broad descriptive statements, factor analysis may be the appropriate technique, but if the goal is to make statements concerning relationships among specific variables of theoretical interest, correlation and regression methods would appear to be more appropriate" (Berry 1971, p. 214)

Ecometrics Following the initial excitement in factorial ecology during the 1960s and 1970s, the practice quickly fell out of vogue and lay mostly dormant through the 1980s and 1990s, presumably in part due to its inability to address methodological critiques. After this lull, however, explorations into the factor structure of communities were revived by Raudenbush and Sampson (1999) in a seminal article introducing a newly minted sub-field they term "Ecometrics." To overcome the problems of factorial ecology in the earlier generation, Raudenbush & Sampson propose several improvements to the methodology. First, they argue that while large scale databases like the US Census contain a variety of useful data, they typically fail to capture many of the most important ecological properties of neighborhoods. Instead, Raudenbush & Sampson advocate the use of item-response models tailored specifically to collect information about community structures.

In conjunction with the proposed data collection devices, they encourage the use of *confirmatory* factor analysis (CFA), as opposed to the exploratory factor analyses (EFA) employed by factor ecologists. Confirmatory factor analysis is a special case of structural equation modeling which, unlike its exploratory cousin, allows social scientists to test a-priori theories about factor relationships by specifying a measurement model that describes how particular variables should load onto designated theoretical latent constructs; in so doing, CFA provides an inferential framework for testing whether the social construct under consideration is supported by the data. In this way ecometrics is a marked departure from factor ecology (and arguably a return to Chicago School ideals); whereas the latter is concerned with exploratory urban research, using diverse datasets to examine emergent factors and developing post-hoc interpretations of them, the former is concerned with deductive research. A theory about *why* certain variables are presumed to load into semantically-meaningful factors is stated formally, arguments justifying this are made, and then statistical tests of fit are performed to interrogate these claims (Mujahid et al. 2007).

Ecometrics is still a fledgling methodology, but it has already been shown capable of developing valid and reliable measures of social constructs like collective efficacy, physical disorder, and social disorder, which have important implications for human behavior (Sampson and Raudenbush 1999; Raudenbush and Sampson 1999; Sampson 2002, 2012b). The predominant barrier to adoption in ecometric research has been the costly requirement of systematic social observation (Sampson and Raudenbush 1999). Part of the push for ecometrics was that large-scale administrative data (e.g. the Census) often lack information about the most important ecological properties of communities, and thus novel (and expensive) data collection strategies are necessary. Recently, however, researchers have attempted to make ecometric research more accessible by incorporating "big data" sources and "virtual audits" (O'Brien et al. 2015; Bader et al. 2015, 2017; Sampson 2017). It seems likely that ecometric analysis will continue to grow in popularity, particularly as additional datasets and new applications materialize. Meanwhile, however, ecometric studies are relatively rare, and the more common practice, by far, is the development of neighborhood classifications and typologies. Instead of factor analysis, these studies employ cluster analysis to identify groups of neighborhoods (i.e. census tracts) whose racial, economic, physical and other attributes are internally homogeneous. This approach is discussed in detail in the following section.

Geodemographics Whereas the early Chicago scholars were interested in *explaining* the relationship between social processes and spatial structure, another pervasive interest has been *describing* the emergence of a particular spatial structure when considering certain socioeconomic, demographic, or behavioral variables. Indeed, "in geographic knowledge discovery the aim is, more often than not, to explore and let spatial patterns surface rather than develop predictive models" (Henriques et al. 2012, p. 218).

This approach represents an important conceptual shift from the factor analytic approaches discussed in Section 2. From a procedural standpoint, both factor analysis and cluster analysis are data reduction techniques sometimes described as "unsupervised machine learning"; whereas factor analysis and geodemographics create composite indices that maintain the greatest amount of information or variance. Thus, factor analysis and geodemographics require a 'speculative synthesis' when determining the meaning of the latent variables. For factor analysis, this requires determining the meaning of loadings; for geodemographics, this requires identifying a meaningful demographic profile for each geodemographic classification (Spielman and Thill 2008). As a result, meaning is synthesized from statistical profiles, instead of generated at the outset through a theory or hypothesis about urban social structure.

Thus, some expressive analysis methods are deductive, seeking to develop theory and test hypotheses about urban ecological processes (e.g. ecometrics). Others focus on inductive analysis, exploring the multitude of ways that urban segregation manifests without specifying the axes of differentiation (e.g. geodemographics). As ecometrics modernizes factorial ecology, so too does geodemographics modernize social area analysis.

In this sense, geodemographics reorient social area analysis away from sociological models of spatial structure to geographic ones. The sociological line of inquiry is arguably about location choice: why do different groups of people come to inhabit discrete parts of the city? It is also concerned with social process: how do spatial contexts (and the social systems that develop within them) influence collective behaviors like altruism or crime? In addressing these questions, urban sociologists wanted to know if the same factor structure emerged in different places and different societies–if so, it would represent anthropological evidence that human social and political behavior is influenced by some kind of natural laws. By contrast, the geographic line of inquiry is arguably about location intelligence: how different are any two neighborhoods, based on the estimated profiles of their residents? What can we learn about cities by studying how residential areas are split by different classifications? In both cases, geography is an *expression* or manifestation of the underlying fundamental process at hand, rather than an object of study itself.

Over the last several decades, geodemographics has become a common avenue of academic study and a lucrative enterprise for private industry market research. Indeed, "the analysis of people by where they live" (Petersen et al. 2011, p. 174) has become the dominant approach for characterizing how socio-spatial structure is expressed in urban areas. Much like BCZ, geodemographic approaches "organize areas into categories sharing similarities across multiple socioeconomic attributes" (Singleton and Spielman 2014, p. 558). The distinguishing difference between SAA and geodemographics is that the latter does not employ factor analysis prior to clustering. Thus, rather than describing the essential components of urbanism, geodemographic classifications are themselves "small area indicators of the social, economic and demographic conditions prevailing in small areas, or 'neighbourhoods'." As statistics themselves, these classifications can flexibly incorporate any kind of urban data (Singleton and Longley 2009b, p. 289). This flexibility means that geodemographic approaches can be tailored to a wide variety of purposes, but also raises the challenge in "substantiating that they reflect real divisions in society, not chance grouping in the data" (Singleton and Spielman 2014, p. 563).

Geodemographic segmentation systems have been applied with success to a wide variety of practical settings including public health (Farr and Evans 2005; Abbas et al. 2009; Petersen et al. 2011), education (Harris et al. 2007; Singleton and Longley 2009a; Singleton et al. 2012), criminal justice (Ashby and Longley 2005), marketing (Dalton and Thatcher 2015), road safety (Brown et al. 1999; Anderson 2010), urban microsimulation (Birkin and Clarke 2012), and several others (Singleton and Spielman 2014). In the realm of public policy, Webber and Burrows (2018) show how the city of Liverpool has been using geodemographics for decades to develop better urban plans, and Batey and Brown (2007) develop a geodemographic method for assessing whether government initiatives are serving adequately their intended spatial targets. In the private sector, meanwhile, geodemographic systems like MOSAIC and ACORN have flourished over the last several decades, enabling marketing and financial service providers to better target customers using geodemographics to model customer behavior (Farr and Webber 2001). Towards this end commercial products have proven enormously powerful and consume voracious amounts of data through partnerships with aggregators like Experian and other financial vendors (Webber and Burrows 2018).

3 Urban Regional Science: Embodying Urban Contexts

All expressive methods seek to analyze how urban space *expresses* social difference, which is done either by identifying the distinct effect neighborhoods have on their inhabitants or by estimating unique classifications/demographic profiles that apply to pre-existing demographic areas. In contrast to this, the *embodied* approaches of urban regional science take urban space as constitutive of social difference. Instead of identifying how (pre-existing) neighborhoods are divided by sociodemographic structures, spatially-coherent neighborhoods are constructed that embody these divides. Alternatively, instead of estimating the effect of context on its inhabitants, the shape of context itself is distilled from its inhabitants. Thus, whereas expressive methods use geography as a medium to express social structure, embodied methods identify coherent geographies latent within sociodemographic structure.

What a "coherent" geography means, though, requires the core analytical concept of regional science, the "region." In urban regional science, a "region" is a spatially-bounded territory that stands in for a conceptually- or mathematically-relevant target of analysis. Since regions are "spatially-bounded," they are usually exclusive (meaning that observations can only be in one region) and exhaustive (all observations are in at least one region). Thus, regions completely partition the urban space under study. Their use (and thus, relevance) depends on the context being studied (Openshaw 1977). With respect to identifying neighborhoods, regionalization methods operationalize (Galster 2001), finding "bundles of spatially based attributes associated with clusters of residences." But, urban regional science approaches focus on more than neighborhoods alone, allowing for meaningful "bundles of spatially based attributes" that pertain to a wide variety of distinct urban locations (residences, but also workplaces, commute paths, leisure spaces, etc.). Focusing exclusively on urban regional science *about neighborhoods*, then, bounding these coherent bundles of spatially based attributes identifies how relevant social processes are *embodied* within urban space. The methods, techniques, & common operational theories used to estimate these boundaries are called "regionalization."

Distinct from "clustering," regionalization requires the partitioning of a map into a finite number of exclusive labellings. Map clustering seeks only to identify unusual regions, even those that are geographically irregular (Kolatch 2001) or do not provide an exhaustive partition.† Thus, clustering is a fully "unsupervised" analytical technique, whereas regionalization is often described as semi-supervised. Generally-speaking, the analyst has a notion of how many regions are desired, geographical conditions the regions should satisfy (such as compactness, convexity, and/or contiguity), and which implicit geographies the detected regions might echo (Duque et al. 2007). However, in nearly all cases, the recovery of an existing neighborhood geography is not the end of regionalization, so it is not a strictly supervised technique.

Beyond the core unifying concept of the embodied "region," regionalization methods have a much wider and diffuse set of applications & techniques. Because regional delineations are strongly dependent on how the process plays out in space, regionalization methods themselves usually do not relate to specific hypotheses about social systems. Instead, regionalization involves a large set of broadly useful methods for partitioning geographies. This can make the literature on regionalization appear more diffuse than the expressive methods discussed previously. However, this diffusiveness is a necessary companion of maturing geographic perspectives (Johnston et al. in press); there are few grand "geographic theories" in the same sense as those considered by the Chicago School, only specific theories about the geography of each social process. In light of this, we present the regionalization in the following section by identifying commonality in both methods & applications. For each case study we discuss, we identify common regionalization strategies, examine shared conceptual entities that regions are used to represent, and describe how these embodied geographies might relate to other studies' geographies.

Regionalization methods

In their review of regionalization algorithms, Duque et al. (2007) identify five criteria used for drawing regions. They describe the various conditions governing how "areas," the fundamental units of observation being grouped, are usually grouped together into the "regions" defining a regionalization. Below, we name and paraphrase the conditions suggested Duque et al. (2007) that regionalizations tend to satisfy:

1. *exclusiveness* - observations are in at most one region
2. *exhaustiveness* - observations are in at least one region
3. *fullness* - each region has more than one observation
4. *disjunction* - each region has a distinctive geographic location and does not overlap or blend into another
5. *optimality* - the regionalization is designed to score well according to a formally-specified objective

Thus, a regionalization algorithm usually provides a full, exclusive-exhaustive partition of a source graph (designed to represent the urban geography under analysis) into many distinctive parts. Taken together, these subgraphs satisfy some target goal or objective; this objective might be explicitly spatial, purely sociodemographic, or may reflect a mixture of any number of component objectives.

Fully-Exclusive Regionalizations: bounding the neighbourhood

Work on the fundamental theory of how best to conduct regionalization analyses, in general, is active and ongoing (Folch and Spielman 2014; Laura et al. 2015; Kim et al. 2016; She et al. 2017). Although regions are sometimes required purely for statistical purposes (Openshaw 1977; Spielman and Folch 2015), regions are often used to model urban residential markets (Royuela and Duque 2013), social communities (Hipp 2010; Hipp et al. 2012, 2013), political communities (Morrill 1976; Guo 2008; Pang et al. 2010; Tam Cho and Liu 2016; Magleby and Mosesson 2018), disease clusters (Assuncao et al. 2006), and transit zones (Guo and Bhat 2007; Li et al. 2014; Chen et al. 2015). ‡

Depending on the social process under study and the frame of analysis, these may be larger or smaller than other common notions of how "large" a neighborhood is from the perspective of the expressive literature (Spielman

† For a recent overview of these cluster detection methods, consult Grubesic et al. (2014).

‡ Beyond the urban, regionalization is common in the identification of ecological zones (Long et al. 2010; Miele et al. 2014; Yuan et al. 2015; Long and Robertson 2018)

and Singleton 2015). Often, these analyses compare the identified data-driven regions to an existing regionalization, identifying how and where the the solutions agree or examining which observations tend to be ill-fitting. This means that many analyses consider the number of observed regions as if it reflects a "known" or true number of admissible regions. This is not a necessary constraint (Duque et al. 2012), however, and the number of admissible (or intelligible) regions has itself be used to analyze volatility in neighborhood dynamics (Rey et al. 2011) or to provide more useful statistical summaries of small-area estimates (Bação et al. 2004; Henriques et al. 2010; Assunçao et al. 2006; Spielman and Folch 2015).

The Fuzzy Urban Region

In geoscience & nature geography, many regionalizations have allowed for classifications which are not strictly disjoint (Bourgault et al. 1992; Leyk and Zimmermann 2007; Long et al. 2010; Yuan et al. 2015). In these cases, it is reasonable to consider the regions being embodied as only partially-identifiable. Ecological or geological zones may reasonably blend smoothly into one another, creating spaces where samples might plausibly fall into more than one cluster/region.

Only some of their bounds, edges, or extents are discernible, mainly where the difference in empirical characteristics between regions is largest.

However, by dint of constraint, classical regionalization methods may *force* these partially-identified regions into being complete exclusive-exhaustive assignments. This is akin to some of the concerns discussed by Spielman and Singleton (2015); uncertainty both about which region a site ought to fall into and uncertainty about the site itself may affect classifications across the board. What Kwan (2012) refers to as the "Uncertain Geographic Context Problem"– this fundamental epistemological uncertainty about the scale and precise hierarchy at which theorized regions affect observations–is an intrinsically difficult representational problem. In general, since one cannot know the "true" regions that individuals find most relevant or most impactful for a given social process (or combination of processes), misspecification of the relevant regions may result in statistical or empirical artifacts; observations may be assigned the incorrect contextual effect, multiple contexts may act jointly and their effects are not identifiable, observations may be mis-assigned and thus bias an existing contextual effect estimate away from its "true" value were the set of all regions known.

Indeed, this is a fundamental concern: Isard (1956) identifies this problem right from the outset of urban regional science. He notes, "[regional scientists] shall probably never be in the position to identify a 'true' set of regions," so they are forced to use a new purpose-driven regionalization for each distinct interrogation. Thus, the "true" context, in Isard's view, was likely to remain uncertain. However, through repeated study, commonalities in the structure of relevant regions would emerge, possibly leading to regions which minimized the extent to which they obscured the social processes they co-constituted. While these are just now coming within reach for advanced statistical studies (Bradley et al. 2017), the extent to which these regions represent intelligible socially-experienced geographies is currently unknown. Thus, while some analyses do aim to critically consider uncertainties and measurement (Harris et al. 2007; Gale and Longley 2013; Singleton et al. 2016; Knaap 2017), practical consideration of the uncertain structure of urban regions in this literature is surprisingly rare given the issue's longstanding theoretical attention.

Beyond uncertainty in classification, the inherent rigidity of assuming regions are *disjoint*, which means that the identified zones may be more separated or distinguished geographically than they may be in theory. As some of the hierarchical and spectral methods note, classifications need not be *strictly* disjoint; indeed, it is often reasonable to think that regions or neighborhoods may have "fuzzy" boundaries. This uncertainty of boundary is distinct from uncertainty in classification or measurement; if regions are useful insofar as they identify a distinct territory, then blending or interlacing assignments at the boundary may denote areas where a single region assignment may not be useful or accurate. There are a few attempts to generalize these concerns in classic regionalization methods, either by considering membership itself as a fuzzy decision (Ambroise et al. 1997; Ambroise and Govaert 1998; Cowpertwait 2011; Hu et al. 2009; Reich and Bondell 2011) or by allowing a component assumption of *disjointedness* to be relaxed (Spielman and Logan 2013; Yuan et al. 2015; Wolf 2018). For instance, in Spielman and Logan (2013), street segments are classified into ethnic categories using historical census data. However, classifications are not exhaustive, in that some streets are not identified as being of any discernible ethnic category. Further, "neighborhoods" are loosely-bounded, allowing for the intermingling of classifications into the same bounded space. However, in these studies, neighborhoods *qua* regions still represent a single zone, albeit less crisply-bounded; in the case of social applications, these zones reflect shared contexts that are experienced by many individuals in a shared socio-spatial urban geography. It is only the assumptions about regional structure—that they must be exclusive, exhaustive, and disjoint—that have been relaxed.

Rejecting the Neighborhood-as-Region There are also outright rejections of exclusiveness, exhaustiveness, or disjunction. In the main, these rejections of neighborhoods as regions are theoretically motivated. They may suggest that only a subset of boundary/transitional areas is well-defined. These bounding approaches focus exclusively on identifying zones of rapid change rather than providing membership into discrete categories (in a processes referred to as "wombling") (Womble 1951; Bocquet-Appel and Bacro 1994; Lu et al. 2007; Dean et al. 2018). Another rejection of classical region assumptions in urban regional science involves the rejection of *shared* context. Here, "egohoods," or spaces of individual/personal experience (Hipp and Boessen 2013), are used instead. These spaces of individual experience tend to overlap significantly, are usually unique for each individual (Spielman and Thill 2008; Spielman and Yoo 2009; Logan et al. 2011; Spielman et al. 2013) although the characteristics of these spaces can change dramatically depending on how large they are (Fowler 2016). Because the egohood is purely theoretical and obtained usually from straightforward computations applied to individuals' locations, "bounding" an egohood (or shared spaces between egohoods) is not conceptually useful in the same manner as for the region. Thus, it is unusual to consider the co-incidence of egohood boundaries, even though this is a common method of analysis when interviewing individuals about what they perceive their neighborhoods to be (Coulton et al. 2001; Campbell et al. 2009; Coulton et al. 2013; Hwang 2016). Thus, egohoods are conceptually and practically distinct from neighborhoods since they (in general) do not pertain to collections of residences (or transit destinations or sales locations, etc).

Dynamics of Urban Spaces

Bounding the neighborhood is, however, only the first step in the study of neighborhood dynamics. Dating at least back to Isard (1956), *regional dynamics*, the changes in shape and structure of urban regions, have been prised over "comparative statics." Isard believed we were "far from developing meaningful dynamical models which would require a constantly-changing regional pattern" (p.21, *ibid*); yet, in some ways, we are still far from this goal. As Rey et al. (2011) note, neighborhoods can change in two ways. Neighborhood *composition*—the sociodemographic profile that characterizes each neighborhood—can change, and their *boundaries*—the spatial extent over which each neighborhood is considered salient—can drift. Most clearly, processes like gentrification and the "back to the city movement" have considered dynamics primarily in the first case; known & bounded urban places change in their economic or racial characteristics. Further, some attention has been paid to drift & disagreement about boundaries in urban spaces (Hwang 2016) though this usually focuses on change in subjective perceptions or individual identification of urban space. In general, most work in dynamics has been *exploratory*, describing and identifying latent patterns in social-spatial structure, rather than *model-based* in Isard's (or even Raudenbush and Sampson (1999)'s ecometric) terminology.

More generally, there are many strategies which aim to characterize the dynamics of neighborhood change in terms of fixed-boundary changes of composition, the drifting boundaries of urban places, or their covariation. An important review of neighborhood change in expressive analytical frameworks is provided by Schwirian (1983), who outlines several broad patterns of change, including invasion-succession and neighborhood life cycle (which he describes as classical models). He suggests that demographic/ecological, socio-cultural/organizational, political-economy, and social movements can change neighborhood characteristics. But, given that these neighborhoods are assumed to express underlying social structure from the outset, any type of social change could modify how social traits are expressed at any point in urban space.

Recent work on modeling the urban neighborhood change tends to focus on historical footprints of specific urban spaces from a holistic perspective. One vibrant strand of work is found in analyzing the sequences using the optimal matching algorithm (Abbott 1995; Gauthier et al. 2010). In these analyses, an initial geodemographic analysis is adopted to segment underlying urban space into specific neighborhood classes (Mikelbank 2011; Delmelle and Thill 2014; Delmelle 2015, 2016; Ling and Delmelle 2016; Delmelle 2017). Then, the historical experience of an urban area can be examined by summarizing the sequences of the identified neighborhood classifications it experiences. An interesting trait of the current work is that it tends to focus on similarity of experience and could be asynchronous; two areas could be considered to have analogous historical demographic trends if they make the same sequence of transitions at different times. It could be very interesting and rewarding in terms of revealing hidden urban dynamics patterns to fully exploit the optimal matching algorithm and its variants widely used in other fields of social science (Studer and Ritschard 2016). On the other hand, some recent research applied measures of income mobility to provide complementary insights in neighborhood change (Modai-Snir and van Ham 2018). More specifically, the total neighborhood change is decomposed into exchange and structural effects similar to the income mobility tradition (van Kerm 2004). Further cross-fertilization between economic and social dynamics could be explored. Other interesting analyses of neighborhood change are viscerally expressive, examining changes in images of urban streets (Hwang 2015; Naik et al. 2017).

Other perspectives aim at modeling and predicting the dynamics of neighborhoods. As Mikelbank (2011) notes, many neighborhoods experience *deja vu*: they may appear to be transformed, acquiring wholly-new characteristics, but instead quickly change into another common and established type of neighborhood. Thus, work examining how the urban space transitions between neighborhood classes over time provides an alternative characterization of neighborhood dynamics (Aaronson 2001; Rey et al. 2011, 2012; Delmelle 2015; Arribas-Bel and Tranos 2018). These studies model the trajectories of neighborhood classifications as a Markov Chain based on which a transition count/probability matrix can be estimated and further analyzed. In this case, trajectories of neighborhood classes are stepwise examined, rather than being analyzed holistically as a sequence. Some classes might be "stable," in that they may tend to retain their classification rather than change classification frequently. Others may be "volatile," in that they may have a high probability of transitioning to other classes. It should be noted that the classic Markov chains model does not take account of potential spatial effects such as spatial autocorrelation and spatial heterogeneity, as well as temporal heterogeneity. Strategies for diagnostics and extensions could borrow from the regional income distribution dynamics literature (Rey 2001; Rey et al. 2016; Kang and Rey 2018) and interpretation should proceed with caution.

In general, developing these kinds of longitudinal analyses of dynamics has historically been frustrated by the lack of available data. Despite long-running historical projects that focus on specific urban places (Sampson 2012a), long-term high-resolution longitudinal sociodemographic data suitable for spatial analysis are expensive to develop and maintain. So, it has only recently been made freely- and openly-available (Logan et al. 2014, e.g.). This concern also applies to urban street-level imagery, and as such, the recent cultivation of these large and extensive datasets has lifted the prospects for longitudinal analysis of neighborhoods broadly.

4 Issues at the frontier

At present, the analysis of neighborhoods occupies a somewhat enigmatic space in urban studies. Despite the longstanding interest, significant work is still necessary to obtain basic theoretical frameworks for some traits of neighborhoods, such as their usual social- or spatial size (Fowler 2016; Talen 2018). The application of geodemographic typologies has proven useful repeatedly in a variety of settings, and their instrumental value is clear. Further, studies in urban regionalization continue to cast doubt on Isard's hope of a single true context that appears and reappears across many regionalizations (Poorthuis 2018). Regardless, however, the development of regionalization methods tends to be largely descriptive, cross-sectional, and driven strongly by non-academic interests.

Despite their practical utility, geodemographic studies are explicitly non-causal expressions of social structure. Thus, in their pursuit of active, responsive visualization of the social tapestry *as is*, geodemographics has done very little to advance scientific understanding of the underlying social and political *processes* by which spatial segmentation arises and the *mechanisms* by which society or individuals enforce & reinforce it over time. Rich historical accounts of geodemographic work Webber and Burrows (2018) and Singleton and Spielman (2014) detail the success geodemographics have enjoyed in practical settings, but the relative dearth of academic geodemographics is conspicuous.

There are a variety of other critiques which might turn researchers away from geodemographic work (Voas and Williamson 2001). But, geodemographics as a discipline is also exceptionally-responsive to academic critique when it does occur For instance, " on the fly" visualisation of the outcome of geodemographic classifications (Singleton and Longley 2009b) allows for response and interactive segmentation and exploratory spatial data analysis. Linking "of conventional social, economic and demographic geographies to patterns of virtual interactions at fine levels of spatial granularity" (Longley 2012) empowers detailed and sensitive insights to be obtained. For many current and planned geodemographic output area classifications, a broad set of stakeholders is consulted directly. Finally, leveraging "the best spatial analysis methods in a problem centered approach" (Singleton and Longley 2009b, p.290) has supplanted one-size-fits-all approaches.

These improvements aside, one of the longstanding criticisms of geodemographics has been that typologies "fail to advance any real understanding of the social processes which cause neighborhoods to change" (Webber and Burrows 2018, p. 130). One possible reason is that geodemographic analyses are contingent, precisely-oriented to a specific objective on which the demographic segmentation is drawn. This makes them difficult to evaluate objectively. Indeed, "classification can only be deemed 'good' or 'poor' when it has been evaluated in terms of the specific purpose for which it is required" (Singleton and Spielman 2014).

This also means that critical decisions about parameterization have no *true* answer; the "best" number of clusters can be identified (Rousseeuw 1987; Frey and Dueck 2007, e.g.), but this may not provide the clearest or most useful contrasts for the visualization of these clusters. There is considerable variability for this parameter in the literature: five neighborhood types have generally emerged in academic analysis (Mikelbank 2004; Delmelle 2015; Foote and Walter 2017; Xiang et al. 2018), but commercial systems have evolved over time, with original systems using 10 low-level categories, later systems leveraging over 40 (Webber and Burrows 2018), and still others with over 100 (Brown and Batey 1994). Further, hierarchical typologies may nest types-within-types, creating arbitrary numbers and configurations of sub-types within types. The rationale for these decisions is often unclear or uncompelling from a theoretical perspective, but can profoundly affect the resulting structure of identified neighborhoods.

Despite broad utility and appeal in the geographical sciences, a major limitation of geodemographics remains that "current geodemographic systems cannot be considered to be explicitly spatial in design, estimation or testing, and that local context requires systematic consideration in geodemographic profiling" (Singleton and Longley 2009b, p. 289). Further, "it is increasingly clear that what is missing is explicitly spatial representation of the accumulated effects of historical and cultural processes" (Longley 2012). In general, this can foster a good kind of "blindness" in the geodemographic classification. Without extensive consideration of deep history & geography urban demographic segmentations, the analyses may hope to discover novel or interesting patterns, unconstrained by legacies of space and time.

Indeed, the lack of explicit geographical constraint is often viewed as enhancing flexibility more than a limitation, since, as Schwirian points out, "social areas" were originally defined such that "the subareas need not be contiguous. Their similarity arises from the social similarity, not the physical proximity of their residents" (Schwirian 1983, p. 86). Geographies that *are* strongly spatially-similar will generate nearly-contiguous geodemographic classifications, and ones that are not, do not. Moving forward, there is a clear opportunity to bridge expressive and embedded perspectives to include formal models of spatial relationships (Rey and Janikas 2005; Rey et al. 2011, 2015), but these (as of yet) are neither common nor simple.

In contrast, explicit models of spatial regionalization tend to still fixate on methodological developments. Typically, these innovations tend to focus on getting regionalizations faster (Laura et al. 2015), with higher quality (Bradley et al. 2017), or with better guarantees on their properties relative to the rest of an unfathomably-large set of possible regionalization (Chikina et al. 2017). These methodological innovations reflect substantial advances in a known-hard methodological problem, but the relative lack of theoretical work on the validity of urban regionalization methods remains here, too. Further, concerns from geodemographics, such as those about arbitrary numbers of neighborhoods or a solution's contingency on a specific purpose or context, still apply in regionalization (Duque et al. 2012). Thus, while regionalization is often presented more strictly as a methodological problem to be solved through methodological means, it is likely that better theory about the social-spatial structure of neighborhoods could benefit both the current cutting-edge embedded and expressive methods of analysis.

Toward an Urban and Regional Ecology

Contemporary approaches for modeling socio-spatial structure, the expressive perspective employing cluster analysis and the embodied perspective of employing regionalization, are successors to long lines of urban social science. While the former is typically practiced by sociologists as a deductive method for understanding the geography of social processes, the latter is typically practiced by geographers as an inductive method to understand the emergent patterning of spatial structure. Unfortunately, these quantitative disciplines have seldom crossed in the literature, to the detriment of urban scholarship writ large. In our view, these approaches are complements, not substitutes: cross-fertilization between empirical researchers would provide more robust insights into the human processes under study in the city. For most cases, this requires better theory-driven development of methods in neighborhood dynamics.

Addressing common concerns about uncertainty in demographic classifications, future work may aim to integrate the classification and prediction models into the same probabilistic framework. This is currently under active development (e.g. Bradley et al. 2017), and will likely only continue to be more attractive as statistical thinking becomes more expressive and "harder" models become more feasible. Integrating the classification and use model ensures that the uncertainty and flux in classification decisions are formally linked to how they are used.

For instance, many of the geodemographic and regionalization methods are conceptualized as estimating "latent" categories of observations. Recent work makes it simpler to chain theories together in *hierarchical* models, which nest models of multiple processes together in a single consistent probabilistic framework. While hierarchical modeling is not new, formal analysis and computation methods make them more feasible now than they ever have been before. Thus, more extensive use of relevant formal statistical theory is required to address concerns about uncertainty in the analysis of neighborhood dynamics.

Second, it is necessary for more explicit attention to be paid to the human theory underlying neighborhood dynamics. For instance, in analogue to ethnographic studies that use participants' notions of neighborhood boundaries to obtain a consensus about neighborhood structures (Coulton et al. 2001; Campbell et al. 2009; Coulton et al. 2013; Hwang 2016), do neighborhoods of wealth comport with neighborhoods of transit or healthcare or race? Regions or geodemographic partitions can be drawn, but to what extent do the discovered *map patterns* coincide? Current geodemographic and regionalization analyses tend to focus on the latent structure of a large amount of social data, reprising the same data reduction timbre from early Chicago School study. However, since data-driven analyses are destined to be contingent on middle-range theory, it may be more illuminating to embrace this narrowness, examining how patterns in regionalizations or geodemographic partitions of different social processes coincide.

Third, it is important to reconcile the theoretical distinctions made between expressive and embedded analyses. Explicitly-spatial strategies to integrate information about configuration and pattern of the input data into geodemographic classification must be embraced, improved, and deployed. Further, strict beliefs about neighborhoods or regions as contiguous zones need to be relaxed in future work in urban regional science. This may involve methods that explicitly relax constraints (Yuan et al. 2015; Wolf 2018) or generalize their meaning (Wu and Murray 2008). While there are sometimes explicit statistical or mathematical challenges to using spatial data directly within classic geodemographic methods, simpler strategies exist and can be remarkably effective (Halleck Vega and Elhorst 2015). It is important to consider these directions in developments to analyze socio-spatial structure.

Finally, the study of true *dynamics* is critical for relevance in the future. What a true dynamical model of joint spatial and social structure might look like is still uncertain, however. Currently, work proceeds by considering the social structure of tract-years, the "flattened" collection of neighborhoods in space and time. When these are linked together in the analysis of transitions or demographic sequences, there is usually *not* a generative model (in the sense of Schwirian (1983)'s descriptive typologies). Regardless of the specific expressive/embedded approach, this lack of generative dynamics makes it difficult to address Longley (2012)'s concerns about how the accumulation of historical experiences can be invisible or obscured in geodemographic analysis. Better models of neighborhoods as dynamic social-spatial systems may yet prove more sensitive to this legacy of experience.

5 Conclusion

The history of neighborhood dynamics is storied, and involved with many central strands of social science throughout the mid-20th century. However, difficult conceptual challenges, identified at the outset of this genre of study, still plague the domain. This persistent disunity centers on the precise mechanisms and actions by which neighborhoods operate in urban space. Through what social process do neighborhoods (as things in themselves) affect their inhabitants? Which "neighborhoods" (or generally, urban regions) are germane to which processes, and do all neighborhoods play a part in that process? What are neighborhoods for? Despite the longstanding interests in bounding their territory, estimating their effects, and describing their dynamics, these questions are still largely approached by examining as broad of social data as possible, attempting to draw sufficiently general geodemographic classifications using a large amount of census data. While this means the study of neighborhood dynamics often considers the intersection of many different social processes, it still remains contingent on "narrow spatial theorizing," Isard (1956, p. 20) or sometimes no theorizing at all.

Thus, in some ways, the same challenges that have existed continue to exist in nearly the same form, although data and computation have improved immensely. Indeed, there have been so many important, recognizable advances in the methodology for understanding the dynamics of neighborhoods, that it is difficult to characterize exactly how extensive these advances are. New data on new human experiences continue to challenge and change the ways in which we bound and describe neighborhoods. New algorithms and mathematical advances improve the quality and precision with which urban structure can be prescriptively divided or descriptive divisions estimated.

However, most of the theory for modeling neighborhoods and their socio-spatial structure derive from the Chicago School, whose pursuit of human ecology began 100 years ago (Abbott 1997, p. 1153). Many new methods still reprise the same theoretical motifs from this earlier work. And while once a towering paradigm in the social sciences, human ecology has lain dormant for many years. As we have shown, contemporary methods provide two distinct views on how neighborhoods are meaningful in cities. Unfortunately, the growth, maturation, and adoption of methods belonging to the embedded and expressive perspectives have apparently accompanied the erection of disciplinary silos and as a result, modern urban ecology is more diffuse and it is difficult to trace the exact moment of its genesis.

Moving forward we call for a new generation of scholarship that bridges the perspectives, leveraging the useful theoretical backing of the Chicago school with the stronger geographies of urban regional science. Today, we argue, there is an emerging generation of urban ecological scholarship that does precisely this. Though sometimes not actively conscious of this lineage, this new urban ecology traces much of its heritage to the Chicago School. Unlike these predecessors, however, the new urban ecology strives to blend urban regional science embeddings with new expressive analyses. It is this new urban ecology, then, that stands poised to provide new perspectives on the fundamental divide in the study of neighborhood dynamics.

References

Aaronson D (2001) Neighborhood Dynamics. *Journal of Urban Economics* 49(1): 1–31. DOI:10.1006/juec.2000.2181. URL http://linkinghub.elsevier.com/retrieve/pii/S0094119000921818.

Abbas J, Ojo A and Orange S (2009) Geodemographics – a tool for health intelligence? *Public Health* 123(1): e35–e39. DOI:10.1016/j.puhe.2008.10.007. URL http://dx.doi.org/10.1016/j.puhe.2008.10.007http://linkinghub.elsevier.com/retrieve/pii/S003335060800276X.

Abbott A (1995) Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology* 21(1): 93–113. DOI:10.1146/annurev.so.21.080195.000521. URL http://www.annualreviews.org/doi/10.1146/annurev.so.21.080195.000521.

Abbott A (1997) Of Time and Space: The Contemporary Relevance of the Chicago School. *Social Forces* 75(4): 1149. DOI:10.2307/2580667. URL http://www.jstor.org/stable/2580667?origin=crossref.

Alonso W (1964) *Location and Land Use: Toward a General Theory of Land Rent*. Cambridge, MA: Harvard University Press. ISBN 978-0674729568.

Ambroise C, Dang M and Govaert G (1997) Clustering of Spatial Data by the EM Algorithm. In: *Geostatistics for Environmental Applications*, volume 9. pp. 493–504. DOI:10.1007/978-94-017-1675-8_40. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=?doi=10.1.1.25.3634http://link.springer.com/10.1007/978-94-017-1675-8{_}40.

Ambroise C and Govaert G (1998) Convergence of an EM-type algorithm for spatial clustering. *Pattern Recognition Letters* 19(10): 919–927. DOI:10.1016/S0167-8655(98)00076-2. URL http://linkinghub.elsevier.com/retrieve/pii/S0167865598000762.

Anderson TK (2010) Using Geodemographics to Measure and Explain Social and Environment Differences in Road Traffic Accident Risk. *Environment and Planning A* 42(9): 2186–2200. DOI:10.1068/a43157. URL http://journals.sagepub.com/doi/10.1068/a43157.

Anderson TR and Bean LL (1961) The Shevky-Bell Social Areas: Confirmation of Results and a Reinterpretation. *Social Forces* 40(2): 119–124. DOI:10.2307/2574289. URL https://academic.oup.com/sf/article-lookup/doi/10.2307/2574289.

Anderson TR and Egeland JA (1961) Spatial Aspects of Social Area Analysis. *American Sociological Review* 26(3): 392. DOI:10.2307/2090666. URL http://www.jstor.org/stable/2090666?origin=crossref.

Arribas-Bel D and Tranos E (2018) Characterizing the spatial structure (s) of cities "on the fly": The space-time calendar. *Geographical Analysis* 50(2): 162–181.

Arsdol MD, Camilleri SF and Schmid CF (1958a) An Application of the Shevky Social Area Indexes to a Model of Urban Society. *Social Forces* 37(1): 26–32. DOI:10.2307/2573775. URL https://academic.oup.com/sf/article-lookup/doi/10.2307/2573775.

Arsdol MDV, Camilleri SF and Schmid CF (1958b) The Generality of Urban Social Area Indexes. *American Sociological Review* 23(3): 277. DOI:10.2307/2089241. URL http://www.jstor.org/stable/2089241?origin=crossref.

Ashby DI and Longley PA (2005) Geocomputation, Geodemographics and Resource Allocation for Local Policing. *Transactions in GIS* 9(1): 53–72. DOI:10.1111/j.1467-9671.2005.00205.x. URL http://doi.wiley.com/10.1111/j.1467-9671.2005.00205.x.

Assuncao R, Costa M, Tavares A and Ferreira S (2006) Fast detection of arbitrarily shaped disease clusters. *Statistics in medicine* 25(5): 723–742.

Assunçao RM, Neves MC, Câmara G and Da Costa Freitas C (2006) Efficient regionalization techniques for socio???economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science* 20(7): 797–811. DOI:10.1080/13658810600665111. URL http://www.tandfonline.com/doi/abs/10.1080/13658810600665111.

Bação F, Lobo V and Painho M (2004) Geo-self-organizing map (Geo-SOM) for building and exploring homogeneous regions. *International Conference on Geographic Information Science* (November): 22—-37. DOI:10.1007/978-3-540-30231-5. URL http://link.springer.com/chapter/10.1007/978-3-540-30231-5{_}2.

Bader M, Mooney SJ, Bennett B and Rundle AG (2017) The Promise, Practicalities, and Perils of Virtually Auditing Neighborhoods Using Google Street View. *The ANNALS of the American Academy of Political and Social Science* 669(1): 18–40. DOI:10.1177/0002716216681488. URL http://journals.sagepub.com/doi/10.1177/0002716216681488.

Bader M, Mooney SJ, Lee YJ, Sheehan D, Neckerman KM, Rundle AG and Teitler JO (2015) Development and deployment of the Computer Assisted Neighborhood Visual Assessment System (CANVAS) to measure health-related neighborhood conditions. *Health & Place* 31: 163–72. DOI:10.1016/j.healthplace.2014.10.012. URL http://www.sciencedirect.com/science/article/pii/S1353829214001580http://linkinghub.elsevier.com/retrieve/pii/S1353829214001580.

Batey P and Brown P (2007) The Spatial Targeting of Urban Policy Initiatives: A Geodemographic Assessment Tool. *Environment and Planning A* 39(11): 2774–2793. DOI:10.1068/a38519. URL http://journals.sagepub.com/doi/10.1068/a38519.

Bell W (1953) The Social Areas of the San Francisco Bay Region. *American Sociological Review* 18(1): 39. DOI:10.2307/2087847. URL http://search.ebscohost.com/login.aspx?direct=true{&}AuthType=ip,uid{&}db=sih{&}AN=12770865{&}site=ehost-live{&}scope=sitehttp://links.jstor.org/sici?sici=0003-1224{%}28195302{%}2918{%}3A1{%}3C39{%}3ATSAOTS0.CO{%}3B2-O{&}origin=crossref.

Bell W (1955) Economic, Family, and Ethnic Status: An Empirical Test. *American Sociological Review* 20(1): 45. DOI:10.2307/2088199. URL http://www.jstor.org/stable/2088199?origin=crossref.

Bell W and Boat (1957) Urban Neighborhoods and Informal Social Relations. *American Journal of Sociology* 62(4): 391–398. DOI:10.2307/2772880. URL http://www.jstor.org/stable/2772880{%}0Ahttp://www.jstor.org.proxy.uchicago.edu/stable/pdfplus/2772880.pdf?acceptTC=true.

Bell W and Force MT (1956) Urban Neighborhood Types and Participation in Formal Associations. *American Sociological Review* 21(1): 25. DOI:10.2307/2089336. URL http://www.jstor.org/stable/2089336?origin=crossref.

Bell W and Greer S (1962) Social Area Analysis and Its Critics. *The Pacific Sociological Review* 5(1): 3–9. DOI:10.2307/1388270. URL http://www.jstor.org/stable/info/10.2307/1388270.

Berry BJL (1971) Introduction: The Logic and Limitations of Comparative Factorial Ecology. *Economic Geography* 47(4): 209. DOI:10.2307/143204. URL http://www.jstor.org/stable/143204?origin=crossref.

Birkin M and Clarke G (2012) The enhancement of spatial microsimulation models using geodemographics. *Annals of Regional Science* 49(2): 515–532. DOI:10.1007/s00168-011-0472-2.

Bocquet-Appel J and Bacro J (1994) Generalized wombling. *Systematic Biology* 43(3): 442–448.

Bourgault G, Marcotte D and Legendre P (1992) The multivariate (co)variogram as a spatial weighting function in classification methods. *Mathematical Geology* 24(5): 463–478. DOI:10.1007/BF00890530. URL http://link.springer.com/10.1007/BF00890530.

Bourne LS and Barber GM (1971) Ecological Patterns of Small Urban Centers in Canada. *Economic Geography* 47(1): 258. DOI:10.2307/143208. URL http://www.jstor.org/stable/143208?origin=crossref.

Bradley JR, Wikle CK and Holan SH (2017) Regionalization of multiscale spatial processes by using a criterion for spatial aggregation error. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79(3): 815–832.

Brindley T and Raine J (1979) Social Area Analysis and Planning Research. *Urban Studies* 16(3): 273–289. DOI:10.1080/

713702552. URL http://journals.sagepub.com/doi/10.1080/713702552.

Brown PJB and Batey P (1994) Design and Construction of a Geodemographic Targeting System: Super Profiles 1994. Technical Report July, The Urban Research and Policy Evaluation Regional Research LaboratoryUniversity of Liverpool, Liverpool. DOI:10.13140/2.1.1162.8162.

Brown PJB, Hirschfield AFG, Merrall S, Bowers KJ and Marsden J (1999) Targeting Community Safety Projects: The Use of Geodemographics and GIS in the Identification of Priority Areas for Action. In: *European Regional Science Association / Regional Science Association International 39th European Congress - ERSA99*, September 1999.

Burdick-Will JA, Ludwig J, Raudenbush SW, Sampson RJ, Sanbonmatsu L and Sharkey PT (2010) Converging Evidence for Neighborhood Effects on Children's Test Scores: An Experimental, Quasi-Experimental, and Observational Comparison. In: *Whither opportunity?: Rising inequality, schools, and children's life chances*. ISBN 978-0-87154-372-1, pp. 255–276.

Burgess EW (2008) The Growth of the City: An Introduction to a Research Project. In: Marzluff JM, Shulenberger E, Endlicher W, Alberti M, Bradley G, Ryan C, Simon U and ZumBrunnen C (eds.) *Urban Ecology*. Boston, MA: Springer US. ISBN 978-0-387-73412-5, pp. 71–78. DOI: 10.1007/978-0-387-73412-5_5. URL https://doi.org/10.1007/978-0-387-73412-5{_}5http://link.springer.com/10.1007/978-0-387-73412-5{_}5.

Campbell E, Henly JR, Elliott DS and Irwin K (2009) Subjective constructions of neighborhood boundaries: lessons from a qualitative study of four neighborhoods. *Journal of Urban Affairs* 31(4): 461–490.

Chen Y, Xu J and Xu M (2015) Finding community structure in spatially constrained complex networks. *International Journal of Geographical Information Science* 29(6): 889–911.

Chetty R, Hendren N and Katz L (2015) The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. Technical Report August, National Bureau of Economic Research, Cambridge, MA. DOI:10.3386/w21156. URL http://www.nber.org/papers/w21156.pdf.

Chetty R, Hendren N, Kline P and Saez E (2014) Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *. *The Quarterly Journal of Economics* 129(4): 1553–1623. DOI:10.1093/qje/qju022. URL https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qju022.

Chikina M, Frieze A and Pegden W (2017) Assessing significance in a markov chain without mixing. *Proceedings of the National Academy of Sciences* 114(11): 2860–2864.

Coulton CJ, Jennings MZ and Chan T (2013) How big is my neighborhood? individual and contextual effects on perceptions of neighborhood scale. *American journal of community psychology* 51(1-2): 140–150.

Coulton CJ, Korbin J, Chan T and Su M (2001) Mapping residents' perceptions of neighborhood boundaries: a methodological note. *American journal of community psychology* 29(2): 371–383.

Cowpertwait PSP (2011) A regionalization method based on a cluster probability model. *Water Resources Research* 47(11). DOI:10.1029/2011WR011084. URL http://doi.wiley.com/10.1029/2011WR011084.

Dalton CM and Thatcher J (2015) Inflated granularity: Spatial "Big Data" and geodemographics. *Big Data & Society* 2(2): 205395171560114. DOI:10.1177/2053951715601144. URL http://journals.sagepub.com/doi/10.1177/2053951715601144.

Dean N, Dong G, Piekut A and Pryce G (2018) Frontiers in residential segregation: Understanding neighbourhood boundaries and their impacts. *Tijdschrift voor economische en sociale geografie* .

Delmelle EC (2015) Five decades of neighborhood classifications and their transitions: A comparison of four US cities, 1970–2010. *Applied Geography* 57: 1–11. DOI:10.1016/j.apgeog.2014.12.002. URL http://linkinghub.elsevier.com/retrieve/pii/S0143622814002860http://dx.doi.org/10.1016/j.apgeog.2014.12.002.

Delmelle EC (2016) Mapping the DNA of urban neighborhoods: Clustering longitudinal sequences of neighborhood socioeconomic change. *Annals of the American Association of Geographers* 106(1): 36–56. DOI:10.1080/00045608.2015.1096188.

Delmelle EC (2017) Differentiating pathways of neighborhood change in 50 U.S. metropolitan areas. *Environment and Planning A* 49(10): 2402–2424. DOI:10.1177/0308518X17722564. URL http://journals.sagepub.com/doi/10.1177/0308518X17722564.

Delmelle EC and Thill JC (2014) Neighborhood quality-of-life dynamics and the great recession: The case of charlotte, North Carolina. *Environment and Planning A* 46(4): 867–884. DOI:10.1068/a46161.

Diez Roux AV (2001) Investigating Neighborhood and Area Effects on Health. *American Journal of Public Health* 91(11): 1783–1789. DOI:10.2105/AJPH.91.11.1783. URL http://ajph.aphapublications.org/doi/10.2105/AJPH.91.11.1783.

Diez Roux AV and Mair C (2010) Neighborhoods and health. *Annals of the New York Academy of Sciences* 1186(1): 125–145. DOI:10.1111/j.1749-6632.2009.05333.x. URL http://search.ebscohost.com/login.aspx?direct=true{&}db=a9h{&}AN=48116272{&}site=ehost-livehttp://doi.wiley.com/10.1111/j.1749-6632.2009.05333.x.

Duque JC, Anselin L and Rey SJ (2012) THE MAX-P-REGIONS PROBLEM*. *Journal of Regional Science* 52(3): 397–419. DOI:10.1111/j.1467-9787.2011.00743.x. URL http://doi.wiley.com/10.1111/j.1467-9787.2011.00743.x.

Duque JC, Ramos R and Suriñach J (2007) Supervised Regionalization Methods: A Survey. *International Regional Science Review* 30(3): 195–220. DOI:10.1177/0160017607301605. URL http://journals.sagepub.com/doi/10.1177/0160017607301605.

Farr M and Evans A (2005) Identifying 'unknown diabetics' using geodemographics and social marketing. *Journal of Direct, Data and Digital Marketing Practice* 7(1): 47–58. DOI:10.1057/palgrave.dddmp.4340504. URL http://link.springer.com/10.1057/palgrave.dddmp.4340504.

Farr M and Webber R (2001) MOSAIC: From an area classification system to individual classification. *Journal of Targeting, Measurement and Analysis for Marketing* 10(1): 55–65. DOI:10.1057/palgrave.jt.5740033. URL http://link.springer.com/10.1057/palgrave.jt.5740033.

Fischer CS (1972) "Urbanism as a Way of Life". *Sociological Methods & Research* 1(2): 187–242. DOI:10.1177/004912417200100203. URL http://journals.sagepub.com/doi/10.1177/004912417200100203.

Folch DC and Spielman SE (2014) Identifying regions based on flexible user-defined constraints. *Int. J. Geogr. Inf. Sci.* 28(1): 164–184.

Foote N and Walter R (2017) Neighborhood and socioeconomic change in emerging megapolitan nodes: tracking shifting social geographies in three rapidly growing United States metropolitan areas, 1980–2010. *Urban Geography* 38(8): 1203–1230. DOI:10.1080/02723638.2016.1217637. URL https://doi.org/10.1080/02723638.2016.1217637https://www.tandfonline.com/doi/full/10.1080/02723638.2016.1217637.

Fowler C (2016) Segregation as a multiscalar phenomenon and its implications for neighborhood-scale research: The case of south Seattle 1990-2010. *Urban geography* 37(1): 1–25.

Frey BJ and Dueck D (2007) Clustering by Passing Messages Between Data Points. *Science* 315(5814): 972–976. DOI:10.1126/science.1136800. URL http://www.ncbi.nlm.nih.gov/pubmed/17218491http://www.sciencemag.org/cgi/doi/10.1126/science.1136800.

Gale CG and Longley PA (2013) Temporal uncertainty in a small area open geodemographic classification. *Transactions in GIS* 17(4): 563–588.

Galster G (2001) On the Nature of Neighbourhood. *Urban Studies* 38(12): 2111–2124. DOI:10.1080/00420980120087072. URL http://journals.sagepub.com/doi/10.1080/00420980120087072.

Galster G (2012) The mechanism (s) of neighbourhood effects: Theory, evidence, and policy implications. In: van Ham M, Manley D, Bailey N, Simpson L and Maclennan D (eds.) *Neighbourhood effects research: New perspectives*. New York: Springer. ISBN 9400723083, pp. 23–56.

Galster G (2017) The Geography of Opportunity 20 Years Later. *Housing Policy Debate* 27(6): 941–943. DOI:10.1080/10511482.2016.1216745. URL https://www.tandfonline.com/doi/full/10.1080/10511482.2016.1216745.

Galster GC and Sharkey PT (2017) Spatial Foundations of Inequality: A Conceptual Model and Empirical Overview. *RSF: The Russell Sage Foundation Journal of the Social Sciences* 3(2): 1. DOI:10.7758/rsf.2017.3.2.01. URL http://www.jstor.org/stable/10.7758/rsf.2017.3.2.01.

Garner CL and Raudenbush SW (1991) Neighborhood Effects on Educational Attainment: A Multilevel Analysis. *Sociology of Education* 64(4): 251. DOI:10.2307/2112706. URL https://www.jstor.org/stable/2112706?origin=crossrefhttp://www.jstor.org/stable/2112706http://www.jstor.org/stable/2112706?origin=crossref.

Gauthier J, Widmer E and Bucher P (2010) Multichannel sequence analysis applied to social science data. *Sociological* : 1–38URL http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9531.2010.01227.x/full.

Greer S (1956) Urbanism Reconsidered: A Comparative Study of Local Areas in a Metropolis. *American Sociological Review* 21(1): 19. DOI:10.2307/2089335. URL http://www.jstor.org/stable/2089335?origin=crossref.

Grubesic TH, Wei R and Murray AT (2014) Spatial Clustering Overview and Comparison: Accuracy, Sensitivity, and Computational Expense. *Annals of the Association of American Geographers* 104(6): 1134–1156. DOI:10.1080/00045608.2014.958389. URL http://www.tandfonline.com/

doi/abs/10.1080/00045608.2014.958389.

Guo D (2008) Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *International Journal of Geographical Information Science* 22(7): 801–823. DOI:10.1080/13658810701674970. URL http://www.tandfonline.com/doi/abs/10.1080/13658810701674970.

Guo JY and Bhat CR (2007) Operationalizing the concept of neighborhood: Application to residential location choice analysis. *Journal of Transport Geography* 15(1): 31–45.

Halleck Vega S and Elhorst JP (2015) The SLX Model. *Journal of Regional Science* 55(3): 339–363.

Harris R, Johnston R and Burgess S (2007) Neighborhoods, Ethnicity and School Choice: Developing a Statistical Framework for Geodemographic Analysis. *Population Research and Policy Review* 26(5-6): 553–579. DOI:10.1007/s11113-007-9042-9. URL http://link.springer.com/10.1007/s11113-007-9042-9.

Hawley AH (1950) *Human Ecology: A Theory of Community Structure*. Ronald Press. ISBN 9780471069386. URL https://books.google.com/books?id={_}GvXAQAACAAJ.

Hawley AH and Duncan OD (1957) Social Area Analysis: A Critical Appraisal. *Land Economics* 33(4): 337. DOI:10.2307/3144311. URL http://www.jstor.org/stable/3144311?origin=crossref.

Henriques R, Bacao F and Lobo V (2010) Spatial Clustering with SOM and GeoSOM: Case Study of Lisbon's Metropolitan Area. In: *2010 Second International Conference on Advanced Geographic Information Systems, Applications, and Services*. IEEE. ISBN 978-1-4244-5809-7, pp. 148–152. DOI:10.1109/GEOProcessing.2010.30. URL http://ieeexplore.ieee.org/document/5437658/.

Henriques R, Bacao F and Lobo V (2012) Exploratory geospatial data analysis using the GeoSOM suite. *Computers, Environment and Urban Systems* 36(3): 218–232. DOI:10.1016/j.compenvurbsys.2011.11.003. URL http://dx.doi.org/10.1016/j.compenvurbsys.2011.11.003http://linkinghub.elsevier.com/retrieve/pii/S0198971511001141.

Herbert D (1967) Social Area Analysis: a British Study. *Urban Studies* 4(1): 41–60. DOI:10.1080/00420986720080031. URL http://journals.sagepub.com/doi/10.1080/00420986720080031.

Hipp JR (2010) Micro-structure in micro-neighborhoods: A new social distance measure, and its effect on individual and aggregated perceptions of crime and disorder. *Social Networks* 32(2): 148–159. DOI:10.1016/j.socnet.2009.11.001. URL http://dx.doi.org/10.1016/j.socnet.2009.11.001.

Hipp JR and Boessen A (2013) Egohoods as waves washing across the city: A new measure of "neighborhoods". *Criminology* 51(2): 287–327.

Hipp JR, Butts CT, Acton R, Nagle NN and Boessen A (2013) Extrapolative simulation of neighborhood networks based on population spatial distribution: Do they predict crime? *Social Networks* 35(4): 614–625. DOI:10.1016/j.socnet.2013.07.002. URL http://dx.doi.org/10.1016/j.socnet.2013.07.002http://linkinghub.elsevier.com/retrieve/pii/S037887331300066X.

Hipp JR, Faris RW and Boessen A (2012) Measuring 'neighborhood': Constructing network neighborhoods. *Social Networks* 34(1): 128–140. DOI:10.1016/j.socnet.2011.05.002. URL http://dx.doi.org/10.1016/j.socnet.2011.05.002.

Hu T, Ouyang J, Qu C and Liu C (2009) Initialization of the Neighborhood EM Algorithm for Spatial Clustering. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 5678 LNAI. ISBN 3642033474, pp. 487–495. DOI:10.1007/978-3-642-03348-3_48. URL http://link.springer.com/10.1007/978-3-642-03348-3{_}48.

Hunter A (1971) The Ecology of Chicago: Persistence and Change, 1930-1960. *American Journal of Sociology* 77(3): 425–444. DOI:10.2307/2776178. URL http://search.ebscohost.com/login.aspx?direct=true{&}AuthType=ip,uid{&}db=sih{&}AN=15404476{&}site=ehost-live{&}scope=site.

Hunter A (1975) The Loss of Community: An Empirical Test Through Replication. *American Sociological Review* 40(5): 537. DOI:10.2307/2094194. URL http://www.jstor.org/stable/2094194{%}0Ahttp://www.jstor.org/stable/2094194?origin=crossrefhttp://www.jstor.org/stable/2094194?origin=crossref.

Hunter AA (1972) Factorial Ecology: A Critique and Some Suggestions. *Demography* 9(1): 107. DOI:10.2307/2060548. URL http://link.springer.com/10.2307/2060548.

Hwang J (2015) Gentrification in changing cities: Immigration, new diversity, and racial inequality in neighborhood renewal. *The Annals of the American Academy of Political and Social Science* 660(1): 319–340.

Hwang J (2016) The Social Construction of a Gentrifying Neighborhood. *Urban Affairs Review* 52(1): 98–128. DOI:10.1177/1078087415570643. URL http://journals.

sagepub.com/doi/10.1177/1078087415570643.

Isard W (1956) Regional science, the concept of region, and regional structure. *Papers in Regional Science* 2(1): 13–26.

Janson Cg (1971) A Preliminary Report on Swedish Urban Spatial Structure. *Economic Geography* 47: 249. DOI:10.2307/143207. URL http://www.jstor.org/stable/143207?origin=crossref.

Johnston R, Harris R, Jones K, Manley D, Wang WW and Wolf L (in press) Quantitative methods i: The world we have lost – or where we started from. *Progress in Human Geography* DOI:10.1177/0309132518774967. URL https://doi.org/10.1177/0309132518774967.

Kang W and Rey SJ (2018) Conditional and joint tests for spatial effects in discrete markov chain models of regional income distribution dynamics. *The Annals of Regional Science* DOI:10.1007/s00168-017-0859-9. URL https://doi.org/10.1007/s00168-017-0859-9.

Kim K, Dean DJ, Kim H and Chun Y (2016) Spatial optimization for regionalization problems with spatial interaction: a heuristic approach. *International Journal of Geographical Information Science* 30(3): 451–473. DOI:10.1080/13658816.2015.1031671. URL http://dx.doi.org/10.1080/13658816.2015.1031671http://www.tandfonline.com/doi/full/10.1080/13658816.2015.1031671.

Knaap E (2017) The Cartography of Opportunity: Spatial Data Science for Equitable Urban Policy. *Housing Policy Debate* 27(6): 913–940. DOI:10.1080/10511482.2017.1331930. URL https://www.tandfonline.com/doi/full/10.1080/10511482.2017.1331930.

Kolatch E (2001) Clustering algorithms for spatial databases: A survey. Technical report, University of Maryland, College Park. DOI:10.1.1.28.1145. URL http://mrl.cecsresearch.org/Resources/papers/Ericasurvey.pdf.

Kwan MP (2012) The uncertain geographic context problem. *Annals of the Association of American Geographers* 102(5): 958–968.

Landay S (1971) The Ecology of Islamic Cities: The Case for the Ethnocity. *Economic Geography* 47(1): 303. DOI:10.2307/143212. URL http://www.jstor.org/stable/143212?origin=crossref.

Laura J, Li W, Rey SJ and Anselin L (2015) Parallelization of a regionalization heuristic in distributed computing platforms – a case study of parallel- p -compact-regions problem. *International Journal of Geographical Information Science* 29(4): 536–555. DOI:10.1080/13658816.2014.987287. URL http://dx.doi.org/10.1080/13658816.2014.987287http://www.tandfonline.com/doi/full/10.1080/13658816.2014.987287.

Leyk S and Zimmermann NE (2007) Improving land change detection based on uncertain survey maps using fuzzy sets. *Landscape Ecology* 22(2): 257–272.

Li W, Church RL and Goodchild MF (2014) An extendable heuristic framework to solve the p-compact-regions problem for urban economic modeling. *Computers, Environment and Urban Systems* 43: 1–13. DOI:10.1016/j.compenvurbsys.2013.10.002. URL http://dx.doi.org/10.1016/j.compenvurbsys.2013.10.002http://linkinghub.elsevier.com/retrieve/pii/S0198971513000872.

Ling C and Delmelle EC (2016) Classifying multidimensional trajectories of neighbourhood change: a self-organizing map and k-means approach. *Annals of GIS* 22(3): 173–186. DOI:10.1080/19475683.2016.1191545. URL http://dx.doi.org/10.1080/19475683.2016.1191545.

Logan JR (1978) Growth, Politics, and the Stratification of Places. *American Journal of Sociology* 84(2): 404–416. DOI:10.2307/2777855. URL http://www.jstor.org.proxy.lib.ohio-state.edu/stable/2777855.

Logan JR, Spielman S, Xu H and Klein PN (2011) Identifying and Bounding Ethnic Neighborhoods. *Urban Geography* 32(3): 334–359. DOI:10.2747/0272-3638.32.3.334. URL http://www.tandfonline.com/doi/abs/10.2747/0272-3638.32.3.334.

Logan JR, Xu Z and Stults BJ (2014) Interpolating us decennial census tract data from as early as 1970 to 2010: A longitudinal tract database. *The Professional Geographer* 66(3): 412–420.

Long J, Nelson T and Wulder M (2010) Regionalization of Landscape Pattern Indices Using Multivariate Cluster Analysis. *Environmental Management* 46(1): 134–142. DOI:10.1007/s00267-010-9510-6. URL http://link.springer.com/10.1007/s00267-010-9510-6.

Long J and Robertson C (2018) Comparing spatial patterns. *Geography Compass* 12(2): e12356.

Longley PA (2012) Geodemographics and the practices of geographic information science. *International Journal of Geographical Information Science* 26(12): 2227–2237. DOI:10.1080/13658816.2012.719623. URL http://www.tandfonline.com/doi/abs/10.1080/13658816.2012.719623.

Lu H, Reilly CS, Banerjee S and Carlin BP (2007) Bayesian areal wombling via adjacency modeling. *Environmental and Ecological Statistics* 14(4): 433–452.

Magleby DB and Mosesson DB (2018) A new approach for developing neutral redistricting plans. *Political Analysis* 26(2):

147–167.

McKenzie RD (1924) The Ecological Approach to the Study of the Human Community. *American Journal of Sociology* 30(3): 287–301. DOI:10.1086/213698. URL http://www.journals.uchicago.edu/doi/10.1086/213698.

Mendenhall R, DeLuca S and Duncan G (2006) Neighborhood resources, racial segregation, and economic mobility: Results from the Gautreaux program. *Social Science Research* 35(4): 892–923. DOI:10.1016/j.ssresearch.2005.06.007. URL http://linkinghub.elsevier.com/retrieve/pii/S0049089X05000293.

Merriman B (2015) Three Conceptions of Spatial Locality in Chicago School Sociology (and Their Significance Today). *American Sociologist* 46(2): 269–287. DOI:10.1007/s12108-014-9239-4.

Miele V, Picard F and Dray S (2014) Spatially constrained clustering of ecological networks. *Methods in Ecology and Evolution* 5(8): 771–779. DOI:10.1111/2041-210X.12208.

Mikelbank BA (2004) A typology of U.S. suburban places. *Housing Policy Debate* 15(4): 935–964. DOI:10.1080/10511482.2004.9521527.

Mikelbank BA (2011) Neighborhood Déjà Vu: Classification in Metropolitan Cleveland, 1970-2000. *Urban Geography* 32(3): 317–333. DOI:10.2747/0272-3638.32.3.317. URL http://www.tandfonline.com/doi/abs/10.2747/0272-3638.32.3.317.

Mills ES (1967) An Aggregative Model of Resource Allocation in a Metropolitan Area. *The American Economic Review* 57(2): 197–210. DOI:10.2307/1821621. URL http://pum.princeton.edu/pumconference/papers.htmlhttp://www.jstor.org/stable/1821621{%}5Cnhttp://www.jstor.org/stable/1821621.

Modai-Snir T and van Ham M (2018) Structural and exchange components in processes of neighbourhood change: A social mobility approach. *Applied Spatial Analysis and Policy* DOI:10.1007/s12061-017-9249-z. URL https://doi.org/10.1007/s12061-017-9249-z.

Morrill RL (1976) Redistriciting revisited. *Annals of the Association of American Geographers* 66(4): 548–553.

Morris FB and Pyle GF (1971) The Social Environment of Rio de Janeiro in 1960. *Economic Geography* 47(1): 286. DOI:10.2307/143210. URL http://www.jstor.org/stable/143210?origin=crossref.

Mujahid MS, Diez Roux AV, Morenoff JD and Raghunathan T (2007) Assessing the Measurement Properties of Neighborhood Scales: From Psychometrics to Ecometrics. *American Journal of Epidemiology* 165(8): 858–867. DOI:10.1093/aje/kwm040. URL https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwm040.

Murdie RA (1969) *Factorial ecology of metropolitan Toronto, 1951-1961: an essay on the social geography of the city*. Number nos. 115-117 in University of Chicago. Department of geography. Research paper. [University of Chicago] Dept. of Geography. URL https://books.google.com/books?id=JjwDAAAAMAAJ.

Muth RF (1969) Cities and housing; the spatial pattern of urban residential land use. *American Journal of Agricultural Economics* 91(1): 19–41. URL http://trid.trb.org/view.aspx?id=545388.

Naik N, Kominers SD, Raskar R, Glaeser EL and Hidalgo CA (2017) Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences* 114(29): 7571–7576. DOI:10.1073/pnas.1619003114. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1619003114.

Newton PW and Johnston RJ (1976) Residential Area Characteristics and Residential Area Homogeneity: Further Thoughts on Extensions to the Factorial Ecology Method. *Environment and Planning A: Economy and Space* 8(5): 543–552. DOI:10.1068/a080543. URL http://journals.sagepub.com/doi/10.1068/a080543.

O'Brien DT, Sampson RJ and Winship C (2015) Ecometrics in the Age of Big Data. *Sociological Methodology* 45(1): 101–147. DOI:10.1177/0081175015576601. URL http://journals.sagepub.com/doi/10.1177/0081175015576601.

Openshaw S (1977) A Geographical Solution to Scale and Aggregation Problems in Region-Building, Partitioning and Spatial Modelling. *Transactions of the Institute of British Geographers* 2(4): 459. DOI:10.2307/622300. URL http://www.jstor.org/stable/622300http://www.jstor.org/stable/622300?origin=crossref.

Palm R and Caruso D (1972) Factor Labelling In Factorial Ecology. *Annals of the Association of American Geographers* 62(1): 122–133. DOI:10.1111/j.1467-8306.1972.tb00848.x.

Pang S, He H, Li Y, Zhou T and Xing K (2010) An approach of redistricting based on simple and compactness. In: *International Conference in Swarm Intelligence*. Springer, pp. 415–424.

Park RE, Burgess EW and McKenzie RD (1925) *The City*. Chicago: Chicago, IL: University of Chicago Press. URL http://www.jstor.org/stable/3004850?origin=crossref.

Parker JBA (1975) The Factorial Ecology of Dublin: A Preliminary Investigation. *Economic and Social Review*

URL http://edepositireland.ie/bitstream/handle/2262/69075/v7n11975{_}3.pdf?sequence=1{&}isAllowed=y.

Perle ED (1979) Variable Mix and Factor Stability in Urban Ecology. *Geographical Analysis* 11(4): 410–414. DOI:10.1111/j.1538-4632.1979.tb00708.x. URL http://doi.wiley.com/10.1111/j.1538-4632.1979.tb00708.x.

Petersen J, Gibin M, Longley P, Mateos P, Atkinson P and Ashby D (2011) Geodemographics as a tool for targeting neighbourhoods in public health campaigns. *Journal of Geographical Systems* 13(2): 173–192. DOI:10.1007/s10109-010-0113-9.

Poorthuis A (2018) How to draw a neighborhood? the potential of big data, regionalization, and community detection for understanding the heterogeneous nature of urban neighborhoods. *Geographical Analysis* 50(2): 182–203.

Raudenbush SW and Sampson RJ (1999) Ecometrics: Toward a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods. *Sociological Methodology* 29(1): 1–41. DOI:10.1111/0081-1750.00059. URL http://onlinelibrary.wiley.com/doi/10.1111/0081-1750.00059/abstract{%}5Cnhttp://doi.wiley.com/10.1111/0081-1750.00059http://journals.sagepub.com/doi/10.1111/0081-1750.00059.

Rees PH (1969) The Factorial Ecology of Calcutta. *American Journal of Sociology* 74(5): 445–491. DOI:10.1086/224681. URL http://www.journals.uchicago.edu/doi/10.1086/224681.

Rees PH (1971) Factorial Ecology: An Extended Definition, Survey, and Critique of the Field. *Economic Geography* 47(4): 220. DOI:10.2307/143205. URL http://www.tandfonline.com/doi/full/10.2753/JEI0021-3624440403http://www.jstor.org/stable/143205?origin=crossref.

Reich BJ and Bondell HD (2011) A Spatial Dirichlet Process Mixture Model for Clustering Population Genetics Data. *Biometrics* 67: 381–390.

Rey S, Anselin L, Li X, Pahle R, Laura J, Li W and Koschinsky J (2015) Open Geospatial Analytics with PySAL. *ISPRS International Journal of Geo-Information* 4(2): 815–836. DOI:10.3390/ijgi4020815. URL http://www.mdpi.com/2220-9964/4/2/815/http://www.mdpi.com/2220-9964/4/2/815.

Rey SJ (2001) Spatial empirics for economic growth and convergence. *Geographical Analysis* 33(3): 195–214.

Rey SJ, Anselin L, Folch DC, Arribas-Bel D, Sastré Gutiérrez ML and Interlante L (2011) Measuring Spatial Dynamics in Metropolitan Areas. *Economic Development Quarterly* 25(1): 54–64. DOI:10.1177/0891242410383414. URL http://journals.sagepub.com/doi/10.1177/0891242410383414.

Rey SJ and Janikas MV (2005) Regional convergence, inequality, and space. *Journal of Economic Geography* 5(2): 155–176. DOI:10.1093/jnlecg/lbh044. URL https://academic.oup.com/joeg/article-lookup/doi/10.1093/jnlecg/lbh044.

Rey SJ, Kang W and Wolf L (2016) The properties of tests for spatial effects in discrete markov chain models of regional income distribution dynamics. *Journal of Geographical Systems* 18(4): 377–398. DOI:10.1007/s10109-016-0234-x. URL http://dx.doi.org/10.1007/s10109-016-0234-x.

Rey SJ, Mack EA and Koschinsky J (2012) Exploratory Space–Time Analysis of Burglary Patterns. *Journal of Quantitative Criminology* 28(3): 509–531. DOI:10.1007/s10940-011-9151-9. URL http://link.springer.com/10.1007/s10940-011-9151-9.

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics* 20: 53–65.

Royuela V and Duque JC (2013) Housi: Heuristic for delimitation of housing submarkets and price homogeneous areas. *Computers, Environment and Urban Systems* 37: 59–69.

Salins PD (1971) Household Location Patterns in American Metropolitan Areas. *Economic Geography* 47(1): 234. DOI:10.2307/143206. URL http://www.jstor.org/stable/143206?origin=crossref.

Sampson RJ (2002) Transcending tradition: New directions in community research, Chicago style. *Criminology* 40(2): 213–230. DOI:10.1111/j.1745-9125.2002.tb00955.x. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.2002.tb00955.x/abstracthttp://doi.wiley.com/10.1111/j.1745-9125.2002.tb00955.x.

Sampson RJ (2012a) *Great American City*. University of Chicago Press. ISBN 9780226055688. DOI:10.7208/chicago/9780226733883.001.0001. URL http://www.press.uchicago.edu/ucp/books/book/chicago/G/bo5514383.htmlhttp://www.asanet.org/journals/CC/MAR13CCFeature.pdfhttp://www.bibliovault.org/BV.landing.epl?ISBN=9780226734569.

Sampson RJ (2012b) *Great American city: Chicago and the enduring neighborhood effect*. University of Chicago Press.

Sampson RJ (2017) Urban sustainability in an age of enduring inequalities: Advancing theory and ecometrics for the 21st-century city. *Proceedings of the National Academy of Sciences* 114(34): 8957–8962. DOI:10.1073/pnas.1614433114. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1614433114.

Sampson RJ, Morenoff JD and Gannon-Rowley T (2002) Assessing "Neighborhood Effects": Social Processes and New Directions in Research. *Annual Review of Sociology* 28(1): 443–478. DOI:10.1146/annurev.soc.28.110601.141114. URL http://www.annualreviews.org/doi/abs/10.1146/annurev.soc.28.110601.141114http://www.annualreviews.org/doi/10.1146/annurev.soc.28.110601.141114.

Sampson RJ and Raudenbush SW (1999) Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods. *American Journal of Sociology* 105(3): 603–651. DOI:10.1086/210356. URL http://www.journals.uchicago.edu/doi/10.1086/210356.

Sampson RJ, Sharkey P and Raudenbush SW (2008) Durable effects of concentrated disadvantage on verbal ability among African-American children. *Proceedings of the National Academy of Sciences* 105(3): 845–852. DOI:10.1073/pnas.0710189104. URL http://www.pnas.org/cgi/doi/10.1073/pnas.0710189104.

Schmid CF (1950) GENERALIZATIONS CONCERNING THE ECOLOGY OF THE AMERICAN CITY. *American Sociological Review* 15(2): 264–281.

Schmid CF, MacCannell EH and van Arsdol Jr MD (1958) THE ECOLOGY OF THE AMERICAN CITY: FURTHER COMPARISON AND VALIDATION OF GENERALIZATIONS. *American Sociological Review* 23(4): 392–401.

Schwirian KP (1983) Models of Neighborhood Change. *Annual Review of Sociology* 9(1): 83–102. DOI:10.1146/annurev.so.09.080183.000503. URL http://www.jstor.org/stable/2946058http://www.annualreviews.org/doi/10.1146/annurev.so.09.080183.000503.

Sharkey P and Elwert F (2011) The Legacy of Disadvantage: Multigenerational Neighborhood Effects on Cognitive Ability. *American Journal of Sociology* 116(6): 1934–81. DOI:10.1086/660009. URL http://www.journals.uchicago.edu/doi/10.1086/660009.

Sharkey P and Faber JW (2014) Where, When, Why, and For Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects. *Annual Review of Sociology* 40(1): 559–579. DOI:10.1146/annurev-soc-071913-043350. URL http://www.annualreviews.org/doi/abs/10.1146/annurev-soc-071913-043350http://www.annualreviews.org/doi/10.1146/annurev-soc-071913-043350.

She B, Duque JC and Ye X (2017) The Network-Max-P-Regions model. *International Journal of Geographical Information Science* 31(5): 962–981. DOI:10.1080/13658816.2016.1252987. URL http://dx.doi.org/10.1080/13658816.2016.1252987https://www.tandfonline.com/doi/full/10.1080/13658816.2016.1252987.

Shevky E and Williams M (1949) *The social areas of Los Angeles: analysis and typology*. Published for the John Randolph Haynes and Dora Haynes Foundation by the University of California Press. ISBN 9780837156378. URL https://books.google.com/books?id=nrUPAQAAMAAJ.

Sinclair R (1967) Von thünen and urban sprawl. *Annals of the Association of American Geographers* 57(1): 72–87. DOI:10.1111/j.1467-8306.1967.tb00591.x. URL http://www.tandfonline.com/doi/abs/10.1111/j.1467-8306.1967.tb00591.x.

Singleton A, Pavlis M and Longley PA (2016) The stability of geodemographic cluster assignments over an intercensal period. *Journal of Geographical Systems* 18(2): 97–123. DOI:10.1007/s10109-016-0226-x.

Singleton AD and Longley PA (2009a) Creating open source geodemographics: Refining a national classification of census output areas for applications in higher education. *Papers in Regional Science* 88(3): 643–666. DOI:10.1111/j.1435-5957.2008.00197.x. URL http://doi.wiley.com/10.1111/j.1435-5957.2008.00197.x.

Singleton AD and Longley PA (2009b) Geodemographics, visualisation, and social networks in applied geography. *Applied Geography* 29(3): 289–298. DOI:10.1016/j.apgeog.2008.10.006. URL http://dx.doi.org/10.1016/j.apgeog.2008.10.006http://linkinghub.elsevier.com/retrieve/pii/S0143622808000726.

Singleton AD and Spielman SE (2014) The Past, Present, and Future of Geodemographic Research in the United States and United Kingdom. *The Professional Geographer* 66(4): 558–567. DOI:10.1080/00330124.2013.848764. URL http://www.tandfonline.com/doi/abs/10.1080/00330124.2013.848764.

Singleton AD, Wilson AG and O'Brien O (2012) Geodemographics and spatial interaction: an integrated model for higher education. *Journal of Geographical Systems* 14(2): 223–241. DOI:10.1007/s10109-010-0141-5. URL http://link.springer.com/10.1007/s10109-010-0141-5.

Spielman SE and Folch DC (2015) Reducing uncertainty in the American Community Survey through data-driven regionalization. *PLoS One* 10(2): 1–21.

Spielman SE and Logan JR (2013) Using high-resolution population data to identify neighborhoods and establish their boundaries. *Annals of the Association of American Geographers* 103(1): 67–84.

Spielman SE and Singleton A (2015) Studying Neighborhoods Using Uncertain Data from the American Community Survey: A Contextual Approach. *Annals of the Association of American Geographers* 105(5): 1003–1025. DOI:10.1080/00045608.2015.1052335. URL http://www.tandfonline.com/doi/full/10.1080/00045608.2015.1052335.

Spielman SE and Thill JC (2008) Social area analysis, data mining, and GIS. *Computers, Environment and Urban Systems* 32(2): 110–122. DOI:10.1016/j.compenvurbsys.2007.11.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0198971507000907.

Spielman SE and Yoo Eh (2009) The spatial dimensions of neighborhood effects. *Social Science & Medicine* 68(6): 1098–1105. DOI:10.1016/j.socscimed.2008.12.048.

Spielman SE, Yoo EH and Linkletter C (2013) Neighborhood Contexts, Health, and Behavior: Understanding the Role of Scale and Residential Sorting. *Environment and Planning B: Planning and Design* 40(3): 489–506. DOI:10.1068/b38007.

Studer M and Ritschard G (2016) What matters in differences between life trajectories: a comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179(2): 481–511. DOI:10.1111/rssa.12125. URL http://dx.doi.org/10.1111/rssa.12125.

Suttles GD (1972) *The Social Construction of Communities*. Studies of urban society. Chicago: University of Chicago Press. ISBN 9780226781891. URL https://books.google.com/books?id=IsqcQgAACAAJ.

Talen E (2018) In support of the unambiguous neighborhood: a proposed size typology. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability* : 1–23.

Tam Cho WK and Liu YY (2016) Toward a talismanic redistricting tool: A computational method for identifying extreme redistricting plans. *Election Law Journal* 15(4): 351–366.

Taylor PJ and Parkes DN (1975) A Kantian View of the City: A Factorial-Ecology Experiment in Space and Time. *Environment and Planning A: Economy and Space* 7(6): 671–688. DOI:10.1068/a070671. URL http://journals.sagepub.com/doi/10.1068/a070671.

Udry JR (1964) Increasing Scale and Spatial Differentiation: New Tests of Two Theories from Shevky and Bell. *Social Forces* 42(4): 403–413. DOI:10.2307/2574983. URL https://academic.oup.com/sf/article-lookup/doi/10.2307/2574983.

Van Arsdol, MD, Camilleri SF and Schmid CF (1961) An Investigation of the Utility of Urban Typology. *The Pacific Sociological Review* 4(1): 26–32. DOI:10.2307/1388484. URL http://www.jstor.org/stable/10.2307/1388484http://www.jstor.org/stable/info/10.2307/1388271http://www.jstor.org/stable/info/10.2307/1388484.

Van Arsdol MD, Camilleri SF and Schmid CF (1962) Further Comments on the Utility of Urban Typology. *The Pacific Sociological Review* 5(1): 9–13. DOI:10.2307/1388271. URL http://www.jstor.org/stable/info/10.2307/1388271.

van Kerm P (2004) What lies behind income mobility? reranking and distributional change in belgium, western germany and the usa. *Economica* 71(282): 223–239. URL http://www.jstor.org/stable/3549096.

Voas D and Williamson P (2001) The diversity of diversity: a critique of geodemographic classification. *Area* 33(1): 63–76. DOI:10.1111/1475-4762.00009. URL http://www.jstor.org/stable/20004125{%}5Cnhttp://www.jstor.org/stable/pdfplus/20004125.pdfhttp://doi.wiley.com/10.1111/1475-4762.00009.

Webber R and Burrows R (2018) *The Predictive Postcode: The Geodemographic Classification of British Society*. SAGE. ISBN 978-1526402349.

Wirth L (1938) Urbanism as a Way of Life. *American Journal of Sociology* 44(1): 1–24. DOI:10.1086/217913. URL http://www.journals.uchicago.edu/doi/10.1086/217913http://journals.sagepub.com/doi/10.1177/004912417200100203.

Wolf LJ (2018) Spatially-Encouraged Spectral Clustering : A Critical Revision of Spatially-Constrained Spectral Clustering.

Womble WH (1951) Differential systematics. *Science* 114(2961): 315–322.

Wu X and Murray AT (2008) A new approach to quantifying spatial contiguity using graph theory and spatial interaction. *International Journal of Geographical Information Science* 22(4): 387–407. DOI:10.1080/13658810701405615.

Xiang L, Stillwell J, Burns L, Heppenstall A and Norman P (2018) A geodemographic classification of sub-districts to identify education inequality in Central Beijing. *Computers, Environment and Urban Systems* 70(February): 59–70. DOI:10.1016/j.compenvurbsys.2018.02.002. URL https://doi.org/10.1016/j.compenvurbsys.2018.02.002http://linkinghub.elsevier.com/retrieve/pii/S0198971517303198.

Yuan S, Tan Pn, Cheruvelil KS, College LB, Collins SM and Soranno PA (2015) Constrained Spectral Clustering for Regionalization: Exploring the Trade-off between Spatial Contiguity and Landscape Homogeneity. In: *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. ISBN 978-1-4673-8272-4, pp. 1–10. DOI:10.1109/DSAA.2015.7344878.