PCCH

The Center for Complexity in Health

Place and Health as Complex Systems:
A Case Study and Empirical Test

KENT STATE

UNIVERSITY

ASHTABULA

PCCH

Proceedings of the Center for Complexity in Health

PCCH is an annual publication designed both to showcase and provide a publication outlet for some of the main avenues of research being conducted in the *Center for Complexity in Health, Robert S. Morrison Health and Science Building, Kent State University at Ashtabula*. These areas include medical professionalism, community health, allostatic load, school systems, medical learning environments and case-based modeling—all explored from a complexity science perspective.

The studies published in the PCCH are generally comprehensive, in-depth explorations of a topic, meant to provide a wider and more complete empirical and theoretical backdrop for the specific studies that scholars involved in the Center for Complexity in Health (CCH) regularly publish in various disciplinary journals. Such an outlet as the PCCH is useful given the conventions (e.g., page constraints and narrowness of focus) typical of most research periodicals, which make it very difficult to publish relatively complete statements on a topic in complex systems terms. While PCCH studies augment, acknowledge and cite CCH work published in other venues, each PCCH study is an original, distinct manuscript. Finally, PCCH studies are peer-reviewed. Prior to publication each study is sent to colleagues for review and criticism to ensure the highest quality of published proceedings possible.

KENT STATE
UNIVERSITY
ASHTABULA

Kent State University   3300 Lake Road West   Ashtabula, OH 44004   Phone: 440-964-3322

Place and Health as Complex Systems: A Case Study and Empirical Test

**Brian Castellani(1) · Rajeev Rajaram(2) · J Galen Buckwalter(3) ·
Michael Ball(4) · Frederic Hafferty(5)**

Abstract Over the last decade, scholars have developed a *complexities of place* (COP) approach to the study of place and health. According to COP, the problem with conventional research is that it lacks effective theories and methods to model the complexities of communities and so forth, given that places exhibit nine essential "complex system" characteristics: they are (1) causally complex, (2) self-organizing and emergent, (3) nodes within a larger network, (4) dynamic and evolving, (5) nonlinear, (6) historical, (7) open-ended with fuzzy boundaries, (8) critically conflicted and negotiated, and (9) agent-based. While promising, the problem with the COP approach, however, is that its definition remains systematically untested and its recommended complexity methods (e.g., network analysis, agent-based modeling) remain underused. The current article therefore conducts an exhaustive test of all nine characteristics and suggested complexity methods. **Method:** To conduct our test we made two important advances: First, we developed and applied the *Definitional Test of Complex Systems* (DTCS) to a case study on sprawl—a "complex systems" problem—to examine, in litmus test fashion, the empirical validity of the COP's 9-characteristic definition. Second, we used the SACS Toolkit, a case-based modeling technique for studying complex system that employs a variety of complexity methods. For our case study we examined a network of 20 communities (located in Summit County, Ohio USA) negatively impacted by sprawl. Our database was partitioned from the *Summit 2010: Quality of Life Project.* **Results:** Overall, the DTCS found the COP's 9-characteristic definition to be empirically valid. The employment of the SACS Toolkit supports also the empirical novelty and utility of complexity methods. Nonetheless, minor issues remain, such as a need to define health and health care in complex systems terms. **Conclusions:** The COP approach seems to hold real empirical promise as a useful way to address many of the challenges that conventional public health research seems unable to solve; in particular, modeling the complex evolution and dynamics of places and addressing the causal interplay between compositional and contextual factors and their impact on community-level health outcomes.

Keywords Place and Health · Community Health · Complex Systems · Case-Based Modeling · Computational Modeling · Agent-Based Modeling · Networks · Sprawl and Health

1 Background: From Composition to Complexity

Over the last decade, the community health science literature has gone through a series of shifts, moving from the study of composition to the *complexities of place.* This shift began in the early 1990s, when a growing network of researchers moved away from the traditional study of compositional variables to the independent contribution that socioeconomic context has on health [21], [19], [22], [45], [53]. In this new research, compositional factors are defined generally as the aggregation of individual-level variables, such as household income, age, ethnicity, educational level or occupation [50]. Contextual factors are defined generally as the geographical, cultural or socioeconomic conditions/opportunities in which people live [45]. Such factors range from air quality to job growth to a community's health care system [50]. In this new research, compositional factors are seen as relatively independent of (orthogonal to) context. Therefore, when measuring the impact of context, these researchers often partial out (control for) individual or household circumstances [45],[50].

(1)Department of Sociology, Kent State University, 3300 Lake Road West, Ashtabula OH, USA BCASTEL3@KENT.EDU
(2)Department of Mathematics, Kent State University, 3300 Lake Road West, Ashtabula OH, USA
(3)Institute for Creative Technologies, University of Southern California, Marina del Rey, CA, USA
(4)Computer Services Kent State University, 3300 Lake Road West, Ashtabula OH, USA
(5)Program in Professionalism and Ethics, Mayo Clinic, Minneapolis, MN, USA

Between the 1990s and the early 2000s, however, another shift took place, as the study of context conceptually widened into the study of place—the umbrella term for research exploring the role that socio-geographical-contextual forces have on health and wellbeing [19]. In these studies, the places explored included communities, counties, cities, neighborhoods, the built environment, social networks and poverty traps, as well as state-level and national-level differences in health outcomes [19], [54], [56].

The shift from the study of context to place reflected an increasing recognition that the initial conceptualization of context, including its distinction from composition, was theoretically and methodologically problematic [5], [19]. [45] summarized this theoretical problem as such: "'Composition' and 'context' are frequently treated as unproblematic and obvious distinctions, and underlying causal models are often implicit" (p. 129). With little to no theoretical rationale, researchers have overemphasized the differences between context and composition, spending almost no time (causally speaking) thinking about how people and contexts go together to form places. In terms of method, the problem is that places and their health are too complex for linear modeling, variable-based analysis, and overly simplistic causal models that treat compositional and contextual factors as mutually exclusive categories. For researchers such as Cummins, et al [19] and others [45],[54], the failure to recognize the limitation of conventional methods explains the lackluster results that studies of context and health regularly achieve. Often researchers find that, after the effects of various compositional variables are partialled out of an equation, contextual factors have little to no effect [19], [50]. According to [19], these weak results, in turn, lead researchers to employ increasingly sophisticated conventional methodologies, but without yielding much improvement in the results, including rather weak partial-correlation coefficients and "relative risk" ratings of less than 2.0 (See [50]).

1.1 Places as Complex Systems

The theoretical and empirical potential of contextual (place) research, combined with the failure of conventional theories and methods to procure this potential, led Dunn and Cummins to guest-edit a special edition of *Social Science and Medicine* (Volume 65, 2007). The purpose of the various articles in this special edition (which ranged from position papers to exploratory empirical studies) was to point toward new theories and methods that do a better job of "placing health research in context" in order to "address the unanswered

questions about the importance of social and geographical context for health variation" [25]. The new theories embraced included a relational approach to place and health (which comes out of geography, structuration theory, informal reciprocity and actor network theory) and complexity science. The new methods included qualitative method, geographical method, cluster analysis, network method, and agent-based modeling. In terms of unanswered questions, the issues all revolved around arriving at a rigorously defined theory of place that can (a) handle the complex (i.e., nonlinear, emergent, self-organizing, multi-level) causal pathways typical of places and (b) explain how these complex pathways impact health. For example, in terms of our reading of Dunn and Cummins' special edition [25], while arguments vary, a definitional theme exists, which can be stated as such: people and places need to be integrated; related, places need to be thought of in holistic or systems terms as complex, emergent entities; furthermore, places need to be seen as functioning at multiple levels of scale; operating with open-ended boundaries; fluid, mobile and evolving; not constrained by traditional notions of space and time; comprised of nonlinear feedback loops and causal pathways; with subjective histories and multiple social meanings; emerging out of the intersection of the micro and macro, the local and global, and agency and structure; and, finally, as nodes in a larger network of places and environmental forces. In short, places need to be treated as complex systems.

Around the same time that Dunn and Cummins published their special edition, a handful of related articles and books (a few of which were published by the authors included in Dunn and Cummins' special edition) made a similar call. The main difference in these publications, however, was that complexity science, both theoretically and methodologically, was front and center in their argument [2],[6], [22], [23], [28], [32], [38]. They also explicitly theorized place (a.k.a. context, schools, communities, cities, counties, countries) as a complex system. Let us explain.

The cornerstone concept of complexity science is the complex system. As Cilliers [17] and others demonstrate (e.g., [4], [14], [39], [47]), the definition of a complex system is not dictionary in form; instead, it is encyclopedic, and for good reasons. Almost every review of complexity science begins by listing the characteristics central to "their" definition of a complex system. Complex systems, it is variously argued, are self-organizing, adaptive, emergent, comprised of a large number of elements, autopoietic, nonlinear, dynamic, network-like in structure, open-ended with fuzzy boundaries, interdependent, agent-based, evolving, chaotic, comprised of feedback loops, historical, nodes within a larger net-

work of systems, environmentally impacted, etc ([10], [14], [17], [47]). These encyclopedic listings are crucial because, in terms of the theoretical rigor of complexity science, the explanatory power of a complex system's definition (qua theory) comes from its characteristics [17]. Together and in isolation, the characteristics of a complex system not only describe what it is; they explain how it works (Bar-Yam 2003, 2005). In terms of community health science, for example, to say that a place is emergent is to describe and explain how that place works—or at least that is the argument complexity scientists make [47]. Each characteristic listed in a definition constitutes therefore a whole domain of inquiry. The concept of emergence, for example, connects to a line of research that extends back to the late 1800s and classical sociology, systems biology and Gestalt psychology [14], as well as, more recently, to systems science and cybernetics [33], and, starting in the 1980s, to the beginnings of complexity science as defined by the Santa Fe Institute [43], [60] and its key scholars such as Holland [34] and Kauffman [37].

In terms of theorizing places as complex systems, the scholars doing this work have offered their own encyclopedic definitions. [28], for example, offers the following definition of place—which he borrows heavily from [17]. As a complex system, places are comprised of a large number of elements, interacting dynamically across networks, with rich, short-range interactions that can have a widespread impact; where each element (agent) is ignorant of the behavior of the place as a whole; were interactions are generally nonlinear (e.g., feedback loops, etc); and where place is emergent, open-ended, self-organizing, operating in a position far from equilibrium and evolving.

Despite the increased emphasis that Gatrell [28] and colleagues put on complexity science, the "rigorously defined theory of place" they articulate is essentially the same as the relational approach. This similarity is due, in large measure, to the latter's usage of complexity science as well, albeit more informally. For example, both Gatrell and the relational approach emphasize the agent-based, nonlinear, open-ended, emergent, "network like" structure of places. In fact, if Dunn and Cummins' special edition is combined with these related articles, an overarching theoretical definition of "place as a complex system" emerges.

As shown in Table 1, this overarching definition (according to our reading of the literature) includes a total of nine major characteristics. Places are defined as case-based, causally complex configurations (e.g., multi- dimensional, multi-level, feedback loops) that self-organize and emerge out of the compositional and contextual factors of which they are comprised; furthermore, they

are nodes within a larger network; they are dynamic and evolving, spatially clustered, nonlinear, historical and phenomenological; and they are spatially and sociologically open-ended, comprised of conflicted and negotiated power struggles and agent-based. An overarching methodological theme emerges as well, which is as follows: variable-based, linear statistics cannot adequately model the complexities of place, leading instead to poorly designed studies (e.g., overly developed hierarchical regressions), false conceptual distinctions (e.g., treating compositional and contextual variables as orthogonal) and incorrect results (e.g., context is marginally important). To correct this problem, new methods, such as those found in complexity science (e.g., computational modeling, network analysis), are needed ([2],[54]).

1.2 The Research Problem

There is, however, a problem with the COP approach and its critique of conventional theory and method. It has yet to be rigorously defined and empirically tested. To date, while some research has been done to validate the empirical utility of certain key characteristics (See [19], [22], [23], [28], [32], [38]), the majority of work has been positional/speculative in nature, drawing its theoretical and empirical support from disparate areas of inquiry done on this or that aspect of the definition, or from fields outside community health. This lack of testing is problematic because, without hard won, nose-to-the-ground, critical data analysis, it is difficult to explain exactly how studying places as complex systems is an improvement over conventional research. Furthermore, researchers uninterested in or skeptical of these ideas can easily dismiss them—which, in many ways, means the perpetuation of conventional research, despite its significant limitations.

1.3 Research Solution

The purpose of the current article is to advance the community health science literature by conducting an exhaustive test of the empirical validity and theoretical utility of the COP approach, as pertains to a case study on *sprawl and community-level health*. The database used for the current study was partitioned from the *Summit 2010: Quality of Life Project* (Summit-QLP)[58]—see Methods for more information. Using this database, the current article conducts an exhaustive diagnostic test of all nine characteristics and suggested complexity methods. It also makes two important advances: it

TABLE 1
Communities as Complex Systems: List of Key Characteristics

As a complex system, communities are:

---

1   Causally complex (i.e., circular causality, feedback loops, concurrent events taking place at multiple levels), such that context and composition are interdependent.

2   Case-based, complex configurations that self-organize and emerge out of the compositional, contextual and health outcomes factors of which they are comprised.

3   Nodes within a larger network of communities and social forces, such that more attention should be given to the position of places relative to each other and to social forces operating at a larger scale.

4   Dynamic and evolving, usually along various trajectories and around different attracting clusters. Given this perspective, use more dynamic terms such as "declining" or "getting better."

5   Nonlinear (e.g., poverty traps).

6   Historical (e.g., institutional memory) and phenomenological (i.e., people have their own subjective, interpretive frames for understanding and participating in their communities).

7   Spatially and sociologically open-ended with fuzzy boundaries.

8   Comprised of conflicted, negotiated power struggles.

9   Agent-based; comprised of a large number of interacting agents; with agents being mobile and evolving. And yet, while agent-based, complex systems are emergent and self-organizing.

---

offers a new diagnostic test for testing the utility of complexity science definitions and a new complexity-science method, called the SACS Toolkit. Let us explain.

1.3.1 Sprawl: A Complex Systems Problem

For the past few years we have been examining the negative impact sprawl has had on the health and wellbeing of a network of 20 communities in Summit County, Ohio—a typical Midwestern county in the United States that has been hit hard by post-industrialization and the outmigration and clustering of affluent people and resources from the cities into the suburbs.

For us, sprawl is a good case study to test the COP approach because it is, by definition, a complex systems problem. Let us explain.

Sprawl is not something that happens to one community or town. Instead, it is a 'complex systems' phenomenon spread out across, and emerging out of the networked evolution of multiple places, across time.

Equally important, sprawl emerges out of the complex interplay between compositional and contextual factors. In terms of composition, sprawl is a series of microscopic behaviors engaged in by a network of individual agents: families, businesses, etc. More concretely, at the micro-level, it is the unplanned out-migration (flow) of relatively low-density development and residential out-migration into the suburban and semi-rural tiers surrounding an urban area. The unplanned nature of sprawl comes from the fact that, like many complex systems, no single force or agent is steering it. Instead, the system is evolving, self-organizing and emerging on its own, the result of a large number of adaptive, self-focused agents, across different communities, interested in upward social mobility. In turn, from the top-down, sprawl also involves the various macroscopic, community-level arrangements in which the above micro-level behaviors take place. These arrangements constitute the various social institutions operating within and across communities, from health care systems to regional economies. Community health science refers to these macroscopic behaviors and structures as contextual factors.

As sprawl evolves, it creates a geographical network of segregation and exclusion where communities become, in many ways, relative islands in terms of resource usage, politics, wellbeing, etc, with movement amongst communities being largely automobile dependent. A significant, macroscopic consequence of this segregation is that the poorer, urban communities in a place like Summit County, which already have unsatisfactory health outcomes, do not improve ([20], [52]); or, worse, fall into what Bowels, Durlaff and Hoff [7] call a poverty trap: a self-reinforcing situation of persistent and intractable poverty.

When we put all of the above together, we arrive at the following definition: sprawl is a self-organizing, largely unsupervised, nonlinear, dynamic, negotiated ordering of multiple places, all networked together, evolving across time, which emerges out of (a) micro-level, agent-based out-migration and low-level development and (b) the interplay of these micro-level compositional forces with the various and different macro-level structural arrangements (contextual forces) existing across, within and between the evolving network of communities being studied. In short, sprawl is a complex systems problem.

The question, however, is if the COP literature is better than convention at defining and modeling this problem–hence, the purpose of the current study. Our argument was that if, in the process of testing the COP approach, this approach arrived at new and novel information about sprawl, then this approach would be deemed theoretically and methodologically valid.

1.3.2 Study Design

The design of our study is a variation on [38]. In their study, Keshavarz et al [38] used a mixed-methods design (qualitative method and document analysis) to "examine the relevance and usefulness of the concept of 'complex adaptive systems' as a framework to better understand ways in which health promoting school interventions could be introduced and sustained" ([38], p. 1467). For their study they combed through the general literature on complexity science to arrive at a working definition of schools as complex systems. Next, they chose their case study, a set of public schools that had "implemented at least one health promoting schools project." Then, to test the empirical validity of their definition they explored, in litmus test fashion, each of its key characteristics to determine if their case study did, indeed, exhibit the characteristics of a complex system.

Like Keshavarz et al [38], as explained above, our study conducts a litmus test of a case study: sprawl and its impact on the macroscopic health outcomes of a network of 20 communities located in Summit County, Ohio USA. However, our study differs in two important ways:

First, our intention is different. Keshavarz et al [38] wanted to see if the data they collected empirically evidenced the characteristics outlined in their definition. From this perspective, they assumed their definition and its characteristics to be reasonably valid. For them, the issue of validity had to do with whether the case study fit their definition. The intention of our study is the opposite: we wanted to see if the characteristics in Table 1, as defined, should be applied to our data, to determine their degree of fit. In other words, our question was: Does the definition fit the case study? For us, this difference in intent highlights the major problem in complexity science research today, as applied to the social sciences: the definitions of complex systems that social scientists use are generally assumed valid; the only question to be answered is "does the data fit?" However, based on the extensive critiques made by Cilliers [17] and others (e.g., [11], [12], [47]), in the social sciences the real challenge is the opposite: determining the validity of the definitions used, as applied to each and every topic of study. Testing definitions is not, however, a matter of epistemology. The critiques listed above acknowledge the theoretical and methodological plurality of complexity science. Instead, testing is a matter of fit. Is the chosen definition (metaphorical or not) empirically valid and theoretically valuable?

Given our difference in intent, for this study we developed and employed the *Definition Test of Complex Systems* (hereafter referred to as the DTCS). The DTCS is a formal test that guides researchers through the process of assessing the empirical validity of defining their topic as a complex system–see method section for more information.

Second, our methodology is different. Haggis's [32] critique of Keshavarz et al's methodology [38] echoes the second challenge that complexity science makes to the conventions of community health science: if places are complex systems, then new methodologies are necessary. Qualitative analysis or statistics alone cannot get the work done: qualitative analysis cannot handle the large, complex, multi-dimensional, multi-level databases typical of complex systems; and statistics cannot adequately model qualitative variables, nonlinearity or systems-level causality.

Given this methodological challenge, for this study we advanced a case-based complexity science approach to studying places and their health. Specifically, we used the SACS Toolkit to conduct our test. The SACS Toolkit is a case-based, computationally grounded, the-

oretically driven, mixed-methods toolkit for modeling complex systems [15].

Based on our two advancements, our study is organized as follows. We begin with method, where we review our dataset, measures, the SACS Toolkit and the specific techniques used for the current study. Next, we turn to our results, where we employ the DTCS and the SACS Toolkit to examine, in litmus test fashion, each of the nine characteristics listed in Table 1, as applied to our case study on sprawl (Summit County), to determine the empirical validity and theoretical utility of defining place and health as a complex system. Our goal will be to see if and how the COP approach results in new and novel information about sprawl and its impact on the evolving macroscopic health outcomes of a network of communities. If it does, we can assert with some degree of confidence that the COP approach has a degree of theoretical and methodological validity. Finally, we end by summarizing the results of our test and suggesting how the DTCS can be used by other researchers.

2 METHODS

2.1 Summit County and its 20 Communities

The case study for this paper is Summit County, Ohio, USA and its 20 communities. Summit County is typical of many Midwestern communities in the United States: it contains a major city (Akron) struggling to survive globalization and the shift from industrialism to post-industrialism, along with first-tier and second-tier suburbs, as well as a few semi-rural communities. In terms of wealth, there are very poor, urban communities with poverty traps (e.g., Southwest Akron); first-tier suburban, middle-class communities (e.g., Cuyahoga Falls); and very affluent suburban communities, such as Hudson—see Map 1. One finds the typical health disparities in this county as well. For example, while the average "years of life lost per death" for affluent Hudson is 10.5; it is 17.4 for Southwest Akron. In short, there is nothing much anomalous in this county in comparison to most of the western places studied in the community health science literature.

Our identification of the 20 communities in Summit County was based on census tracts data. Summit County is comprised of 121 census tracts. Public health researchers in Summit County created the analytical boundaries of these 20 communities by clustering census tract data according to identifiable communities, cities, towns, neighborhoods and ethnic groupings. Tracts were also clustered to maintain demographic homogeneity within the 20 communities. For more infor-

mation, see the Healthy Summit 2000 Health Indicators Summary Report [58].

2.2 Database

To conduct our tests we partitioned a database from the *Summit 2010: Quality of Life Project* (Summit-QLP). The Summit-QLP is a website that houses twenty years' worth of information on the health and wellbeing of Summit County. All of the data are in the form of PDF reports, providing detailed statistical and qualitative information on Summit County and its 20 communities. To obtain our measures, we combed through the statistical reports, entering data for the 20 communities into a database for 1990 and 2000—the two points in time during which Summit-QLP collected data. We used both time points in our case study because, as we identified above, sprawl is about the evolution of places, and so longitudinal data is important. Similarly, the ultimate goal of complexity science is to study how complex systems change over time. Given the constraints of data collection, however, some of the health outcomes data only represent one point in time—see Table 3 for explanation.
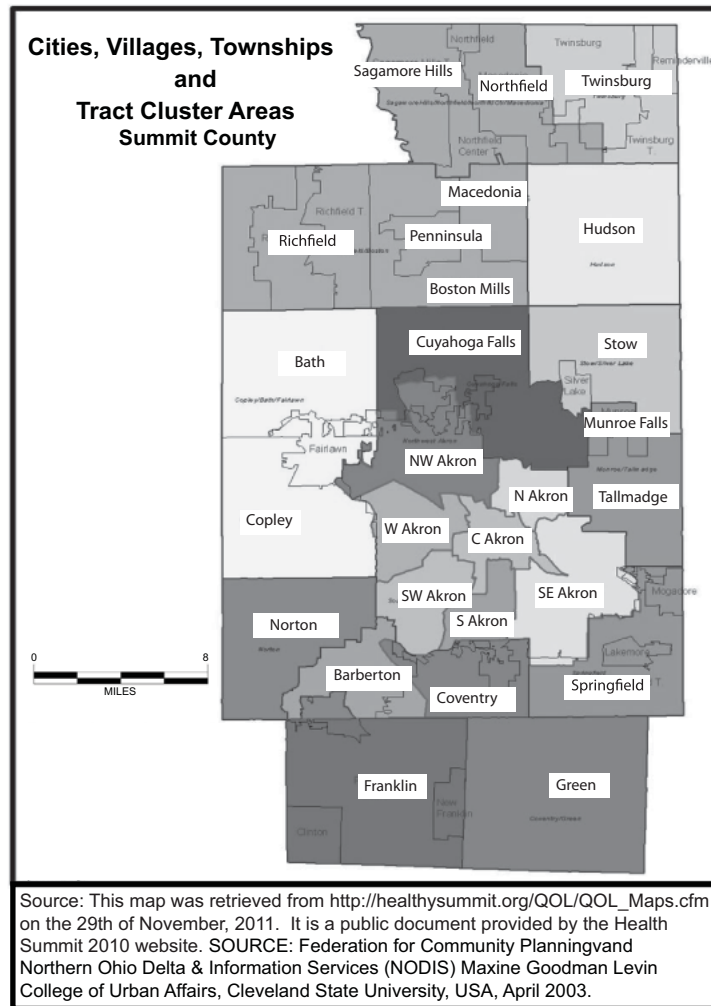
2.3 Measures

As shown in Table 3, the measures used in our study were chosen because they are characteristically used to study sprawl and to operationalize compositional and contextual factors and community-level health outcomes.

Compositional Factors: Two of the most commonly used measures of composition are median household income and ethnicity [45]. When combined, they are also important measures of sprawl: household income and ethnicity allow us to track the micro-level, agent-based, residential migration patterns of affluence and segregation across a network of communities—See Table 3.

Contextual Factors: The argument against sprawl is that it often results in: (a) the economic collapse of urban communities and (b) the segregation of the contextual factors (social institutions) necessary for upward social mobility, such as good schools and a thriving local economy. We therefore sought to include in our database a variety of contextual measures that we could use to track the negative and positive impact sprawl has on communities—in particular, does sprawl lead to a Schelling-like geographical network of segregated affluence [9], with very poor communities stuck in poverty traps. As shown in Table 3, our contextual measures,

Map 1:
Summit County and its 20 Communities

Cities, Villages, Townships and Tract Cluster Areas Summit County

Source: This map was retrieved from http://healthysummit.org/QOL/QOL_Maps.cfm on the 29th of November, 2011. It is a public document provided by the Health Summit 2010 website. SOURCE: Federation for Community Planningvand Northern Ohio Delta & Information Services (NODIS) Maxine Goodman Levin College of Urban Affairs, Cleveland State University, USA, April 2003.

expressed as rates, address a variety of key social institutions, including economy (job growth, civilian labor force, poverty and unemployment), housing (mortgage/rent to income ratio), education (high school completion) and health care (health insurance and public assistance).

Health Outcomes: Sprawl and the community-level segregation of wellbeing are linked to a variety of health outcomes in the literature, in particular early warning measures (e.g., birth weight), adult health measures (e.g., hypertension), mental health (e.g., stress and wellbeing) and mortality (e.g., [8], [26]). For our study, to track the impact of sprawl on community-level health, we employed the following measures: 1st trimester care, childhood immunizations (early life); teen pregnancies (adolescent health); child and elder abuse (mental health); and years of life lost per death (adult health and mortality) —See Table 3.

2.4 The Definitional Test of Complex Systems

Created for the current study, the DTCS is our attempt at an exhaustive tool for determining the extent to which a complex system's definition fits a topic. The DTCS is not, however, a standardized instrument. As such, we have not normed or validated it. Instead, it is a conceptual tool meant to move scholars toward empirically-driven, synthetic definitions of complex systems. To do so, the DTCS walks scholars through a nine-question, four-step process of review, method, analysis, and results—see Table 2.

Following Gatrell, the DTCS does not seek to determine if a particular case fits a definition; instead, it seeks to determine if a definition fits a particular case. As Gatrell explains [28], the challenge in the current literature is not whether places are complex systems; as it would be hard to prove them otherwise. Instead, the

TABLE 2
Definitional Test of Complex Systems (DTCS)

STEP 1: Literature Review and Formulation of the Definition

	QUESTION SET 1: What definition of a complex social system will be used?
	1. What is the definition?
	a. For example, is the definition dictionary in form or encyclopedic?
	b. What are its key characteristics?
	2. Where does the definition come from?
	a. For example, is the definition currently used in the field, or is it a new definition?
	3. What are the definition's epistemological assumptions?
	a. For example, is it postmodern, critical realist, naïve realist, constructionist, etc?
	4. What is the theoretical basis for the definition?
	a. For example, is the definition meant to be metaphorical, literal or prescriptive or some combination?
	5. Does the definition or any of its key characteristics seem to be empirically or theoretically problematic?
	a. For example, are there examples in the literature where usage of the definition led to (i) poorly designed studies, (ii) faulty empirical results, or (iii) flawed or unclear theoretical conclusions?

STEP 2: Methods

	QUESTION SET 2: How will the definition be operationalized and tested?
	6. How will the current test be conducted?
	a. For example, what measures will be used?
	b. What case study will be used?
	c. What analytic techniques will be used for the test?

STEP 3: Run Test

STEP 4: Determine Results

	QUESTION SET 3: What conclusions about the validity and value of the definition were determined?
	7. Did the test suggest that the definition is empirically valid?
	8. Did the test suggest that the definition is theoretically valuable?
	9. In terms of the DTCS's negative hypothesis:
	a. Did the definition or any of its key characteristics lead the test to faulty empirical results?
	b. Did the definition or any of its key characteristics lead the test to flawed theoretical conclusions?
	c. Does the definition obey Occam's razor; or is it a lot of work for little empirical or theoretical yield?

question is: how do we define the complexity of a place? And, does such a definition yield new insights? Given this focus, Question 9 of the DTCS functions as its negative test, focusing on three related issues: the degree to which a definition (a) is being forced or incorrectly used; (b) is not a real empirical improvement over conventional theory or method; or (c) leads to incorrect results or to ideas already known by another name.

Scholars can modify or further validate the DTCS to examine its further utility. Let us briefly review the steps of the DTCS:

STEP 1: To answer the DTCS's initial five questions, researchers must comb through their topic's literature to determine if and how it has been theorized as a complex system. If such a literature does exist, the goal is to organize the chosen definition of a complex system

into its set of key characteristics: self-organizing, path dependent, nonlinear, agent-based, etc. For example, if our review of the community health science literature, we identified nine characteristics. If no such literature exists, or if the researchers choose to examine a different definition, they must explain how and why they chose their particular definition and its set of characteristics, including addressing epistemological issues related to translating or transporting the definition from one field to another.

STEP 2: Next, to answer the DTCS's sixth question, researchers must decide how they will define and measure a definition and its key characteristics. For example, does the literature conceptualize nonlinearity in metaphorical or literal terms? And, if measured literally, how will nonlinearity be operationalized? Once these decisions are made, researchers must decide which methods to use. As we have already highlighted, choosing a method is no easy task. So, social scientists are faced with a major challenge: the DTCS requires them to test the validity of their definitions of a complex system, but such testing necessitate them to use new methods, which many are not equipped to use. It is because of this challenge that, for the current project, we employed the SACS Toolkit, which we discuss next. First, however, we need to address the final two steps of the DTCS.

STEP 3: Once questions 1 through 6 have been answered, the next step is to actually conduct the test. The goal here is to evaluate the empirical validity of each of a definition's characteristics, along with the definition as a whole. In other words, along with determining the validity of each characteristic, it must be determined if the characteristics fit together. Having made that point, we recognize that not all complexity theories (particularly metaphorical ones) seek to provide comprehensive definitions; opting instead to outline the conditions and challenges, for example, that educational administrators face when coming to terms with the complexity of their organizations [48]. Nonetheless, regardless of the definition used, its criteria need to be met.

STEP 4: Finally, with the analysis complete, researchers need to make their final assessment: in terms of the negative test found in question 9 and the null hypothesis of the DTCS, to what extent, and in what ways is (or is not) the chosen definition, along with its list of characteristics, empirically valid and theoretically valuable?

2.5 The SACS Toolkit

To advance the methodological theme of the COP approach, we employed the SACS Toolkit: a case-based, computationally-grounded, mixed-methods toolkit for modeling complex systems [**?**]. The SACS Toolkit was specifically designed to overcome the limitations that conventional methods have in modeling complex systems—which, as Haggis [32] explains, was a major problem with Keshavarz et al.'s [38] study.

The SACS Toolkit is a variation on Byrne's [12] general premise regarding the link between cases and complex systems. For the SACS Toolkit, case-based modeling is the study of complex systems as a set of k-dimensional vectors (cases), which researchers compare and contrast, and then condense and cluster (using computational methods) to create a low-dimensional model of a complex system's topography and dynamics across time/space, while preserving the complexity of the system studied. Such a case-based, computational approach allows the SACS Toolkit to analyze a wide range of highly complex and dynamic data, including the large databases typical in community health science [16]. The other strength of the SACS Toolkit is its assembled nature [16]. While grounded in a defined mathematical framework, it is methodologically open-ended and therefore adaptable and amenable, allowing researchers to employ and bring together a wide variety of computational, mathematical, historical, qualitative and statistical methods. Researchers can even develop and modify the SACS Toolkit for their own purposes. Here we provide a quick, mathematically-grounded overview of the SACS Toolkit. However, for a more in-depth mathematical review, see Castellani and Rajaram [16]; and for an in-depth qualitative review, see Castellani and Hafferty [15].

The SACS Toolkit is comprised of three main components: a theoretical blueprint for studying complex systems (*social complexity theory*); an algorithm for modeling complex systems from the ground up (*assemblage*); and a recommend list of system-friendly modeling techniques (*case-based toolset*).

For detailed information on how to employ the SACS Toolkit, see Castellani and Hafferty [15] and Castellani and Rajaram [16]. Here we provide a brief overview of the modeling algorithm.

Following its theoretical framework, *social complexity theory*, the database the SACS Toolkit assembles for S is comprised of two types of variables: those that make up the complex system of study—which the SACS Toolkit refers to collectively as the web of social practices W—and those regarded as environmental forces E. In the current study, the compositional factors, con-

TABLE 3
Variables Analyzed for the 20 Communities in the Summit County Database

Compositional Factors	• Population 65 years of age of older[1] • % White Population[1] (Defined as number of persons identifying themselves as "White" in response to the 1990 US Census or "White Alone" in response to the 2000 US Census) • % African-American Population[1] (Defined as the number of persons identifying themselves as "Black or African-American" in response to the 1990 US Census or "Black or African-American Alone" in response to the 2000 US Census) • Median Household Income[1]
Contextual Factors	• Overall Poverty[1] (Defined as the number of persons living "below the poverty level" as defined by the U.S. Census) • Public Assistance[1] (Defined as the number of households receive public assistance as defined by the U.S. Census) • Persons 25+ Years with High School Diploma[1] • Net Job Growth[3] (Defined as the number of jobs in 2000 minus the number of jobs in 1990. • Unemployment Rate[1] (Defined as unemployed civilian labor force) • Housing affordability[1] (Defined as the percentage of households where mortgage/rent is greater than 30% of the household income) • No Health Care Coverage[4] (An estimate of the number of individuals with no health care coverage based upon a statewide survey (Behavior Risk Factor Surveillance Survey – Centers for Disease Control and Prevention)
Health Outcomes	• No First Trimester Prenatal Care[4] (Defined as the number of births occurring to mothers from 1995 to and including 1998 for which no prenatal care was received during the first three months of the pregnancy) • Teen Birth Rate[4] (Defined as the number of births occurring between 1995-1998 to mothers 15 to and including 17 years of age) • Childhood Immunization Rate[5] (Defined as the percentage of children with a complete immunization series 4:3:1 by their second birthday based on the kindergarten retrospective study) • Child Abuse/Neglect[6] (Defined as the number of referrals resulting in assessment per 1,000 childre under 18 years of age) • Elder Abuse/Neglect[7] (Defined as the number of referrals received by the Department of Jobs and Family Services for abuse, exploitation, or neglect) • Years of Potential Life Lost per Death[5] (Defined as the sum of the differences between the age at death and the life expectancy at age of death for each death occurring between 1990-1998 due to all causes divided by the number of deaths due to all causes within the census tract cluster borders where those borders are defined by United States Census Bureau census tracts)

Data Sources: (1) United States Census Bureau 1990 and 2000 Decennial Censuses; (2) Ohio Department of Education; (3) NODIS; (4) Akron City Health Department, Office of Epidemiology; (5) Ohio Department of Health; (6) Children's Services Board; (7) Summit County Department of Jobs and Family Service.

textual factors and health outcomes shown in Table 3 form the W for Summit County. In turn, urban sprawl (See $C7$ Results Section) is the E for our study. Together, this set of W and E form the vector configuration for each case.

It therefore follows that, because S consists of n cases $\{c_i\}_{i=1}^n$, and each case c_i has a vector configuration of k-dimensions, it is natural, quantitatively speaking, to represent S, at its most basic, in the form of a data matrix D as follows:

$$D = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} x_{11} & \dots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nk} \end{bmatrix}. \tag{1}$$

In the notation above, the n rows in D represent the set of cases $\{c_i\}$ in S, and the k columns represent

the measurements on some finite partition $\cup_{i=1}^p O_i$ of W and E—as defined in the equation below, which we have written for W:

(a) $O_i \cap O_j = \emptyset \ \forall i \neq j.$ (2)

(b) $\cup_{i=1}^p O_i = W.$

The same definition of the partition of W applies to E to describe a single environmental force or a collection of forces.

Based on a data mining of D—done using the techniques in the case-based toolset—the model that the SACS Toolkit creates is called the *network of attracting clusters* ($\mathcal{N}$). $\mathcal{N}$ is the actual model the SACS Toolkit creates. It is a simplification of the database studied, and is comprised of three types of maps. *Cluster Maps*: Given that the SACS Toolkit studies S as a set of

cases $\{c_i\}$, it is necessary, at some point, particularly in the case of large databases, to identify and map the most common vector configurations in S. An effective method for accomplishing this task is cluster analysis (See [12]). The results of such analyses are the cluster maps for $\mathcal{N}$. In the current study, Table 5 and Figure 2 are examples of such maps. *Network Maps*: In turn, network maps compress S to model the most important relationships (ties, links, etc) and interactions that exist amongst its cases $\{c_i\}$, particularly as they relate to the most common vector configurations in S. In the current study, Figure 4 is an example of such a map. *K-dimensional Maps*: The purpose of the k-dimensional maps is to understand how the variables (W, E) comprising the vector configurations for $\{c_i\}$ influence the structure and dynamics of S. In the current study, examples of such maps include Table 4 and figure 3.

With its set of maps constructed, the next step in the assemblage algorithm is to explore how the maps for $\mathcal{N}$ inform one another, with the goal of arriving at a well-developed, albeit simplified, low-dimensional model of S for one discrete moment in time/space. Figure 1 is one example of what such a model looks like (for one discrete moment across time/space) for our study of Summit County.

Constructing a discrete model of $\mathcal{N}$, however, is not where the SACS Toolkit necessarily ends the assemblage process. For the SACS Toolkit, cases $\{c_i\}$ are ultimately dynamic and evolving. The SACS Toolkit therefore treats cases $\{c_i\}$ as discrete dynamical systems $c_i(j)$, where j denotes the time instant t_j. Thus, the ultimate goal of the SACS Toolkit is to model the case-based structure and dynamics of S as it evolves across time/space. To do so, the SACS Toolkit uses its case-based toolset to generate an $\mathcal{N}$ for each moment across time/space studied. The result of these repeated assemblages of $\mathcal{N}$ is a series of discrete, simplified, low-dimensional models of S that have both dynamical and topographical features. As a final step in the assemblage algorithm, the SACS Toolkit assembles this series of models to form its final case-based model of S. In the current study we were unable to delve into any great detail about the evolution of Summit County across time, as the purpose of our study was to examine the validity of defining places as complex systems. However, tests 2, 4, 6, 7 and 9 deal (in varying degrees) with the issue of dynamics across time/space.

2.5.1 Case-Based Toolset

As stated in the introduction, the main strength of the SACS Toolkit is its assembled nature, which allows it to draw upon and consolidate a wide variety of method-ological techniques to engage in the model building process. In terms of the current study, the following methods were used. To generate our cluster maps (e.g., Table 5 and Figure 2) we used k-means cluster analysis (abbreviated k-means) and the Self-Organizing Map (SOM) [41]. For k-means, we used the statistical software package SPSS. For the SOM we used the SOM Toolbox (www.cis.hut.fi/projects/somtoolbox/download/). The SOM Toolbox runs as a function package in the MATLAB computing environment. We generated our network map (Figure 4) using the freeware Pajek (http://vlado.fmf.uni-lj.si/pub/networks/pajek/). Our k dimensional maps (e.g., figures 3, 4, 6, 7 and 8) were generated using SPSS, the SOM, and agent-based modeling. We built and ran our agent-based model (called Summit-Sim) with NetLogo (ccl.northwestern.edu/netlogo/). Because readers may not know much about these techniques, we provide a brief review below.

SOM and K-Means: Two techniques useful for clustering databases, studying the vector configurations of these cases, and then modeling the evolution of this network of clusters across time are k-means cluster analysis and the Kohonen Self-Organizing Map (SOM) [41]. K-means is a partitional (as opposed to hierarchical), iterative clustering technique that seeks a single, simultaneous clustering solution for some proximity matrix [35]. K-means is also a form of unsupervised learning: unlike classification techniques, the cluster member of a case is not known prior to analysis. The SOM is part of the distributed artificial neural network literature [41]. In this literature, the SOM serves a specific function: mapping high-dimensional data onto a smaller, two-dimensional space, while preserving, as much as possible, the complex, non-obvious patterns of relationships amongst this data [41]. The SOM's strength is its capacity to generate rich, visually intuitive clusters.

The ultimate strength of these two techniques is that they work well together [42]. For example, k-means and the SOM have a similar approach to data compression and clustering. Both can be viewed as vector quantization techniques, insomuch as they cluster cases by searching for a simpler set of reference vectors, with each case in $\{c_i\}$ being positioned near its most similar reference vector. For k-means, the reference vector is a centroid, which represents the average for all the cases in a cluster. For the SOM, the reference vector is an actual point, a neuron, which represents the weighted average of the vector configurations clustering around it. From here, however, they differ. But, it is their differences that make them work so well together.

K-means is useful because it requires the number of centroids to be identified ahead of time, based largely on theoretical rationales. Such an approach is important

Figure 1

Example of the Final Map Created by the SACS Toolkit for Current Case Study
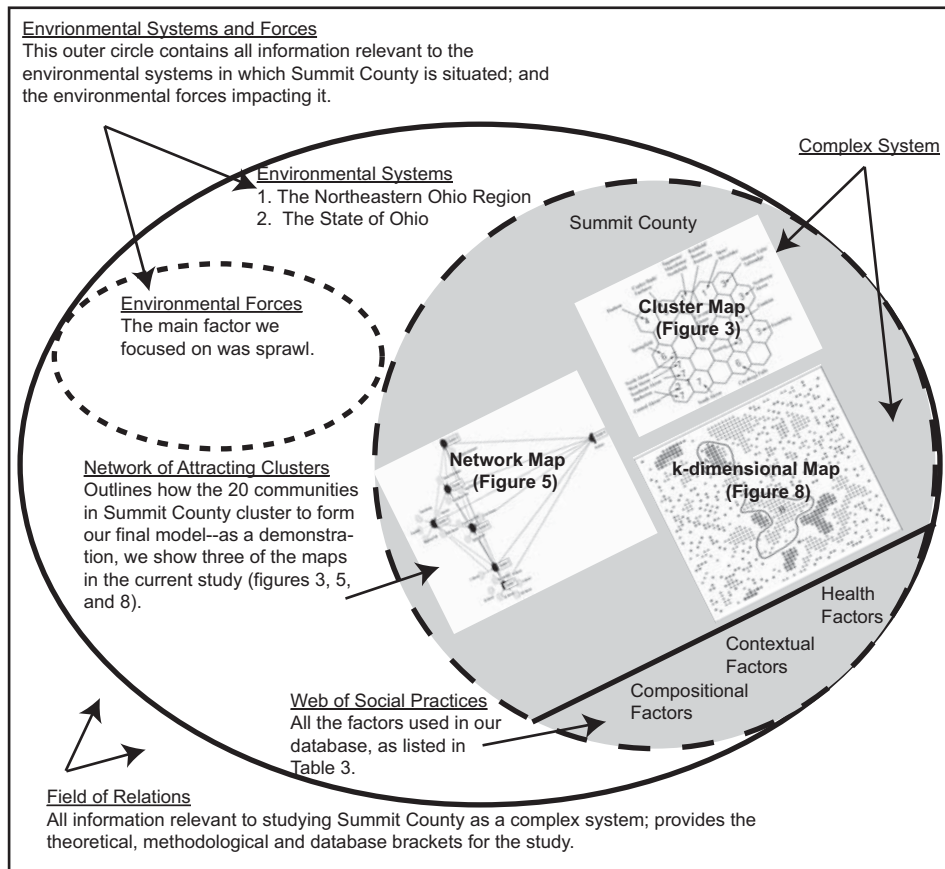
for case-based modeling because it requires that the selection of clusters be somehow theoretically driven— *researchers should have some sense of what they are looking for, based on a preliminary study and comparison of the cases $\{c_i\}$ in S.

In turn, the SOM functions as an effective method of validating k-means because the set of reference vectors (neurons) it settles upon is not predetermined. If, therefore, the SOM arrives at a solution similar to the k-means, it provides an effective method of corroboration. The SOM is also useful because it graphs its reference vectors and the cases $\{c_i\}$ surrounding them as neurons on a two-dimensional, topographical map, called the U-matrix. On the U-matrix, the reference vectors most like one another are graphically positioned as nearby neighbors, with the most unlike reference vectors (neurons) being placed the furthest apart. The U-matrix therefore provides a visually intuitive, low-dimensional map of the original high-dimensional database being studied.

The rationale for using these two techniques to model the evolution of a network of communities across time/

space is as follows: If, following case-based modeling, complex systems typically contain a large number of cases, and if the vector configurations for each case $\{c_i\}$ in a complex system S generally share common profiles (both in terms of proximity and adjacency), then an effective method of longitudinally modeling S, according to the SACS Toolkit, is to cluster it. Clustering is effective because it allows for the identification, mapping and analysis of the most common vector configurations in S for each discrete moment in time/space. These discrete network of attracting clusters can then be treated and mapped as the trajectories (attractor points) for a network of communities across time/space.

Agent-based Modeling: To test Characteristic 9 (C9)– the idea that places are agent-based. Agent-based modeling is a bottom-up approach to simulating complex systems. It is based on the viewpoint that many social outcomes emerge from the micro-level interactions amongst a heterogeneous set of rule-following agents, as they take place across time [29]. The theoretical rationale for using agent-based modeling in the current study is as follows: if sprawl is a complex systems phe-

nomenon that emerges out of the interplay between compositional and contextual factors, then agent-based modeling is necessary, as conventional methods cannot model such an interaction across time.

Summit-Sim: The model built for the current study is called Summit-Sim—See Figure 6. Summit-Sim is a rule-based, multi-agent, discrete-event model with a 51X51 lattice structure, upon which a randomly distributed set of upwardly-mobile agents migrate to find their ideal residential location. Summit-Sim was designed to explore the link between residential migration patterns and health outcomes, based on our empirical study of Summit County and its 20 communities.

For the current study, a simplified version of Summit-Sim was used. Readers can run or downloaded (including code) at cch.ashtabula.kent.edu/summitsim.html). This simple version looks at only one aspect of sprawl: how the residential migration patterns of upwardly mobile agents influence community-level health outcomes. Later developments of this model will incorporate the interplay between residential migration and macroscopic factors such as schooling and job growth. See results section for more details on our theoretical focus.

Figure 6 provides a snapshot of the dashboard for Summit-Sim. On the dashboard are three types of information: the world of Summit-Sim (in the middle); Controls (on the left); and Output Charts (on the right).

Three types of agents inhabit the world of Summit-Sim: rich agents (squares), middle-class agents (triangles) and poor agents (circles)—the number of which is determined by the three population sliders on the left-side of the Summit-Sim dashboard. To test C9 we used several factors from our study of Summit County to determine the population of agents for our simulation: differences in household income (circa 1999/2000); education; the capacity to change residential location; and work type. Poor Agents represent those agents in Summit County who are financially struggling. In 2000, the median household income in Summit County was 42,000. Using Census data, poor agents represent roughly 58 percent of Summit County (with a household income range of 0 to 49,000). Poor agents are associate-level educated or less, with work ranging from low-level white collar work to unemployment. Despite differences, all poor agents have difficulty changing residential location given their financial situation. Middle class agents (31 percent of Summit County) make between 41,000 and 99,000, have some college education or higher, work skilled-blue collar, white-collar or professional jobs, and are moderately able to change residential location. Affluent agents (11 percent of Summit population) represent those households in Summit County making 100,000 or more, who are generally col-lege educated, have professional-class jobs or lucrative blue-collar jobs, and are able to change easily residential location.

In terms of controls, three rules govern the discrete migration behavior of Summit-Sim agents. These rules are Preference, Preference-Degree, and Mobility.
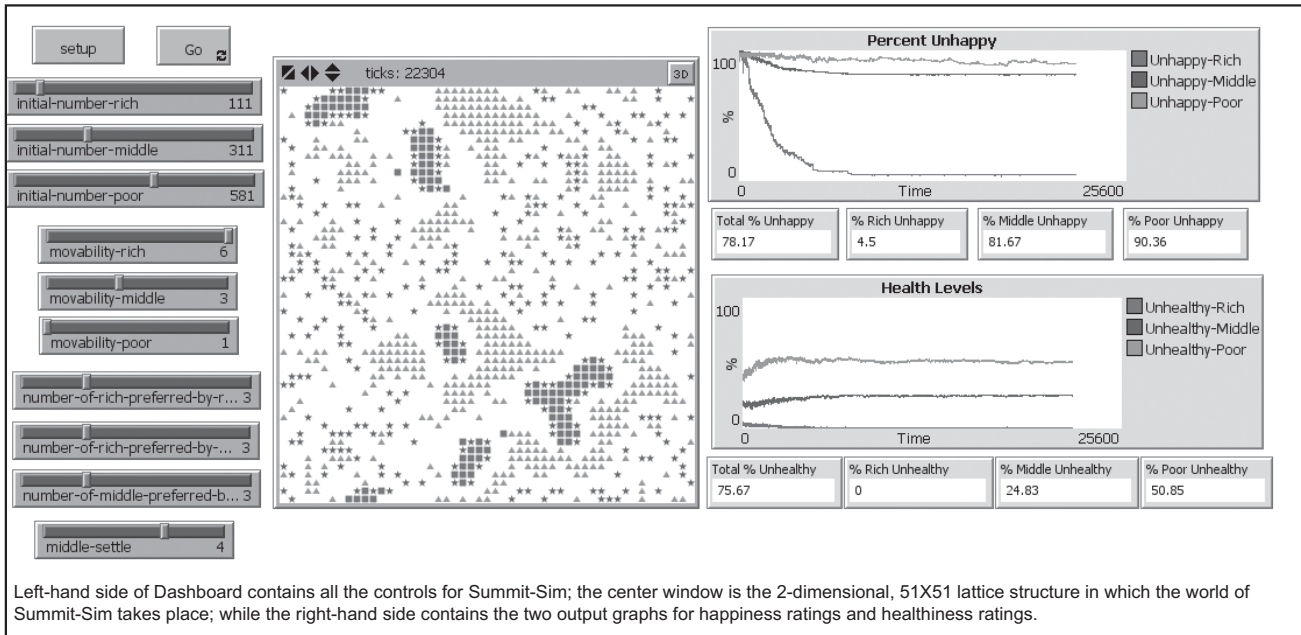
Preference is a modification of Schelling's well-known segregation rule (1978). Unlike the original Schelling model, however, wherein agents seek to live near their own kind, Preference concerns sprawling, upwardly mobile agents migrating to live near agents of a similar or higher status. While sprawl produces segregated neighborhood, it is not necessarily about agents migrating to live near similar agents. Sprawl is about agents migrating to live in better neighborhoods. For rich agents, 'better' means neighborhoods with more rich agents. For middle agents, 'better' means living near more rich agents, or at least lots of middle-class agents. For poor agents, 'better' means living near middle-class agents, if they can. Following this logic, in Summit-Sim, at each discrete point in time, (a) rich agents seek to live near rich agents; (b) middle-class agents seek to live near rich agents; if they cannot, they seek to live near other middle-class agents; if they find themselves in a neighborhood with 4 or more middle-agents, they stay; and (c) poor agents seek to live near middle-class agents; if they cannot, they stay where they are.

Preference-Degree determines the number of higher status agents around which agents prefer to live. In a 2-D lattice structure, 'neighbors' is defined as the total number of spaces (squares) available around an individual agent, which range from 0 to 8. In Summit-Sim, preference ranges from 1 to 3. Our more exhaustive tests find that, if preference is set beyond 3, the model is unable to settle.

Movability is the capacity for an agent to migrate to the neighborhood in which they ultimately desire to live. Ranging from 1 to 6, mobility is defined as the number of spaces an agent can move per iteration. Following our empirical analysis of Summit County, we set the movability of poor agents at 1, primarily because it is very difficult for these agents to buy homes, sell homes or rent a more expensive apartment in order to move. We set the movability for middle-class agents at 3, because they are moderately able to sell their homes or buy a new home or rent a more expensive apartment in order to change location. And, we set the movability at 6 for rich agents because they can move with little effort.

Summit-Sim is comprised of two major charts. The first chart is an unhappiness rating. At each iteration agents are asked if they are happy. Happiness is defined as living in the type of neighborhood they seek.

FIGURE 6:
Snapshot of the Netlogo Dashboard for Summit-Sim

Left-hand side of Dashboard contains all the controls for Summit-Sim; the center window is the 2-dimensional, 51X51 lattice structure in which the world of Summit-Sim takes place; while the right-hand side contains the two output graphs for happiness ratings and healthiness ratings.

The Unhappiness Chart maps the percentage of affluent, middle-class and struggling agents unhappy after each iteration.

The second chart is a healthiness rating. The COP literature generally has shown that residential segregation tends to produce health inequalities insomuch as the less affluent individuals there are in a community, the worse its health outcomes, due in large measure to the complex interplay between compositional factors, such as household income and the health of local institutions, such as schools [7], [20], [52]. Following this argument, we used a rough context-based indicator of community-health. First, we began each simulation of our model with the health of all agents (poor, middle class and rich) being equal. If, however, once the model was started, the 9X9 region in which an agent was living had three or more rich agents, they were considered healthy. At the aggregate level, we were then able to express the healthiness of our three groups as a percentage: overall health, followed by percentages for rich, middle and poor agents.

3 RESULTS

In this section, we will proceed to test each of the COP's nine characteristics, one at a time, in litmus test fashion, to examine the theoretical and methodological utility of the COP approach, as applied to the topic of sprawl. For each characteristic, we summarize what the

COP literature has to say about it, followed by a discussion of the methods we used to test the characteristics and then our results.

3.1 1. Causality is Complex

In terms of studying sprawl and health, a major issue is modeling the complex causal relationship between compositional and contextual factors. We therefore begin with Characteristic 1. The theme of $C1$ takes two forms. In its positive form, the main point of $C1$, as Cummins et al [19] state, is that "research in place and health should avoid the false dualism of context and composition by recognizing that there is a mutual reinforcing and reciprocal relationship between people and place" (p. 1825). In its critical form, the main point of $C1$ is that the "tight interrelationships between individual [composition] and context are not easy to capture in quantitative studies [47]. This is partly why some researchers have adopted important alternative methodological strategies such as qualitative techniques" ([19], p. 1829).

To test $C1$, we used linear modeling to see if we could (according to convention) parse the independent contribution that the compositional and contextual factors listed in Table 3 have with two of our study's health outcomes: Years of Potential Life Lost per Death (YLL) and Teen Birth Rate (TBR). Two notes: (1) the first outcome was chosen because it is a widely used global

measure of community-level health; the second was chosen randomly, so as to report on more than one outcome; however, results similar to those reported below were found when we examined the other outcomes; (2) we analyzed a single time frame for our health outcomes because they are single measures for multiple years, YLL (1990 to 1998) and TBR (1995 to 1998)— See Table 3 for more information.

For our linear modeling we used zero-order correlation first, and then hierarchical regression. We ran zero-order first because it shows the bivariate (pairwise) correlation between each factor and our two outcomes, ignoring the statistical influence that other factors have on the relationship. We ran hierarchical regression second because such analyses demonstrate the "independent" relationship that compositional and contextual factors have with our health outcomes, after the effects of the other factors have been statistically removed from the equation. If, after running hierarchical regression, either set of factors (compositional or contextual) have roughly the same relationship with our two health outcomes (after controlling for the other set of factors) it would support the hypothesis that compositional and contextual factors are independent. Table 4 shows our results: no such "independent" relationship was found. Let us explain.

Column 1 in Table 4 lists the zero-order correlations between our compositional and contextual factors and our two health outcomes. Three things stand out in Column 1. First, YLL and TBR correlate significantly with almost every factor. Second, whatever 'direction of relationship' a compositional or contextual factor had with one health outcome, it had with the other. For example, unemployment in 1990 correlates positively and significantly with both TBR and YLL. Third, several compositional and contextual factors were highly correlated with our two health outcomes. For example, 11 of the 22 correlations were at .75 or better.

Our hierarchical analysis, shown in Column 2, however, reveals a different picture. The only factor that had any remaining impact on YLL was the percentage of the population that was non-Hispanic Caucasian (partial correlation coefficient = -.672; p = .008). No contextual variables (shown in Column 2 as ns, not significant) retained significance when predicting YLL. Conversely, the only factor that had any remaining impact on TBR was the percentage of unemployed (partial correlation coefficient = .702; p = .001); and no compositional variables remained significant when controlling for contextual variables.

Complex linear models rely on the ability to observe predicted patterns. In the current example we see strong zero-order associations, but the associations do not maintain predicted patterns when statistical control is exerted on theoretically relevant variables. This inconsistency may, as the positive form of $C1$ suggests, reflect an inaccurate pattern of predicted associations; that is, compositional and contextual variables are not independent of each other. More likely, given the high level of multicollinearity between our factors, it is likely that our results reflect statistical anomalies. Either way, linear modeling does not provide much information.

3.2 2. Places Are Emergent, Self-Organizing Configurations

In terms of its theme, $C2$ begins where $C1$ ends. If linear statistics cannot adequately model the relationship between health outcomes and compositional and contextual factors, how should researchers model communities? According to $C2$, communities should be modeled as case-based, complex configurations that emerge out of the self-organizing interactions amongst a set of compositional and contextual factors and their related health outcomes.

By emergence, these scholars mean that a community's resulting configuration is such that (a) the whole is more than the sum of its compositional, contextual and health outcomes parts, and (b) one cannot understand this whole through reductionism; the community must be understood as a system [34]. For example, as Curtis and Riva state: "Complexity theory also anticipates that health systems are dynamic and have an inbuilt capacity to organize and reorganize themselves constantly (emergence and re-emergence of human diseases being an illustration)" ([22], p. 2)

As Gatrell [28] explains, by self-organization they mean that the configuration (order) that emerges out of the intersection of a set of compositional and contextual variables and their related health outcomes is: (a) self-sustaining in the face of environmental pressures; (b) not the direct result of any *a priori* design on the part of community, etc; (c) not strictly determined or controlled by any one internal or external supervisor or force; and (d) is more than just the processes of internal feedback loops that can be explained via a linear model ([17], pp. 90-93).

As Blackman [6] explains, by case-based (as discussed in our methods section), they mean two things. First, they mean that places are cases. For these scholars, places are cases in the idiographic sense that each place constitutes a unique configuration of compositional and contextual factors. As Rihoux and Ragin explain [51], a place is a "complex combination of properties, a specific 'whole' that should not be lost or obscured in the course of the analysis—this is a holistic

TABLE 4

Conventional Analysis of Compositional Factors, Contextual Factors and Health Outcomes

		COLUMN 1 Zero-Order Correlations		COLUMN 2 Hierarchical Regression	
		Years of Life Lost Per Death	Teen Birth Rate	Years of Life Lost Per Death	Teen Birth Rate
	Variable				
Compositional Factors	1990 % pop (65yrs&older)	.048* (.840)**	0.435 (.056)	ns	ns
	% Non-Hispanic Caucasian in 1990	-.541 (.014)	-.781 (.000)	-.672*** (.008)	ns
	% African-American in 1990	.551 (.012)	.768 (.000)	ns	ns
	1990 household income	-.794 (.000)	-.814 (.000)	ns	ns
Contextual Factors	% overall poverty 1990	.636 (.003)	.926 (.000)	ns	ns
	% public help 1990	.682 (.001)	.944 (.000)	ns	ns
	% no high school 25yrs+ in 1990	.752 (.000)	.809 (.000)	ns	ns
	Job Growth (1993 to 2000)	-.204 (.387)	-.493 (.027)	ns	ns
	% unemployed 1990	.671 (.001)	.942 (.000)	ns	.702 (.001)
	% of households mortgage/rent is <30% of income	.532 (.016)	.912 (.000)	ns	ns
	% no health care coverage	.684 (.001)	.939 (.000)	ns	Ns

COLUMN 1 provides zero-order, pairwise correlations for all compositional and contextual factors listed in Table 3 with two health outcomes: years of life lost per death and Teen Birth Rate. In this column, (*) is the correlation coefficient; and (**) is its two-tailed, significance level.

COLUMN 2 provides the results of our hierarchical analysis of the "independent" relationships all compositional and contextual factors listed in Table 3 with two health outcomes: years of life lost per death and Teen Birth Rate. In this column (ns) is a non significant partial correlation coefficient; *** is a significant partial correlation coefficient for a two-tailed significance level.

perspective" (p. 6). Second, they mean that, given this holistic view, the study of place cannot be variable-based. Instead, it needs to be case-based, where places are treated as holistic composites of a set of interacting variables and their changes over time. As Blackman [6] states, "complexity theory focuses on cases as empirical and actual domains. The interest is in the 'states' of these cases; not so much how and why variables change but how and why cases change" (p. 31).

Our analysis of Summit County using the SACS Toolkit suggests that $C2$ is an empirically valid and theoretically valuable way to (a) understand the complex causality of places and (b) address the limitations of using conventional method to determine such a relationship.

To test $C2$, we employed two case-based comparative techniques: the self-organizing map algorithm and k-means cluster analysis [15]. As explained in our methods section, we employed these techniques for two reasons. First, they are empirically-driven, iterative techniques designed to explore the configurations that emerge out of the clustered self-organization of a given set of cases and the variables upon which these cases are based— in fact the SOM is literally named the self-organizing map. Second, given the instability and inconsistency often associated with cluster analysis, these techniques corroborate one another [3], [42], [24].

We had two goals for our test of $C2$. Using the SACS Toolkit, our first was to determine if there was any order in the idea that communities are complex configurations that emerge out of the self-organizing interactions amongst a set of compositional and contextual factors and their related health outcomes. To do so, we treated the communities in Summit County as 20 sep-

arate cases, each representing a different configuration of the compositional, contextual and health outcome variables in our study—see Table 3. We began with k-means cluster analysis. We started with k-means because it allowed us to assign our 20 cases to a fixed number of clusters so we could explore different cluster solutions. We also used k-means because it creates single-rank clusters.

For our analyses we used all 17 variables listed in Table 3. We explored normalizing household income because its range was greater than the other variables, but it did not improve the results. We also ran our equation altering variable entry to control for any ordering effect. Finally, we ran the k-means with different fixed cluster solutions. The final solution, shown in Table 5 was a seven cluster solution. We settled on this solution because it separated the cases well without lumping them into any one cluster. Also, it fit with our expert knowledge of Summit County: as suggested by the literature, one effective way to manage the instability of cluster analysis is to have (if possible) prior knowledge about how one's cases should cluster [1].

Here is our breakdown of the clusters. First, there are the affluent suburban communities, which include Hudson (Cluster 4) and Copley Bath and Fairlawn (Cluster 5). Of the two clusters, Hudson is the richest and significantly differs from all other 19 communities, particularly in terms of health outcomes. Next, there are the middle class suburban communities, which include Stow/ Silverlake, Northfield/Macedonia/Sagamore, and Richfield/Peninsula (all in cluster 1). These communities are followed by the middle to working class suburbs and semi-rural areas, which include Springfield, Coventry/Green and Cuyahoga Falls (Cluster 6) and Twinsburg, Northwest Akron, Munroe Falls/Tallmadge, Norton and Franklin (Cluster 3). Of the two, Cluster 3 has a slightly higher average household income and a larger African-American community. Finally, there are the poor inner-city communities, which include all of the communities in the city of Akron (except northwest Akron) as well as the city of Barberton—all of these communities are in Clusters 2 and 7. Cluster 2 has one community, the poorest in Summit County, Central Akron.

With our initial cluster solutions determined, we proceeded to corroborate our k-means with the SOM. The SOM functions as an effective method for validating k-means because the set of reference vectors (neurons) it settles upon is unsupervised. If, therefore, the SOM arrives at a solution similar to the k-means, it provides an effective method of corroboration. The SOM is also useful because, as Figure 2 shows, it spatially graphs its reference vectors (similar to k-means cen-

troids) and the cases (c_i) surrounding them onto a variety of n-dimensional surfaces. The 2-dimensional grid shown in Figure 2b is called a u-matrix, onto which the SOM clustered our 20 communities—the numbers listed on Figure 2b represent the k-means cluster number for each of the 20 communities. Communities distant from one another in Figure 2b are less alike than those closer to one another. As an addition, Figure 2a is a topographical (3-dimensional) u-matrix. On this u-matrix, gray-scale changes indicate conceptual hills and valleys: the lighter the polygon, the greater the conceptual distance (hill) between cases.

Looking at Figure 2, the cluster solution arrived at by the SOM is very similar to the k-means solution. One can see that the richest communities (clusters 4 and 5) are close to each other, with cluster 4 located in the farthest upper left corner—this is Hudson, the richest community. Moving along in clockwise fashion, one finds clusters 1 and then 3 (the next most affluent communities) followed by clusters 6 and 7, and then finally Cluster 2. All of the poor communities in Akron are in the lower left corner of the u-matrix, with the poorest community (Central Akron) residing in the farthest lower left corner. (As a side note: later, when we look at how the 17 variables in our study are distributed across the u-matrix—Figure 3 below—the position of the 20 communities relative to one another will make more sense.) The only real difference between the SOM and the k-means is Cluster 6, which the SOM distributed more widely than the k-means. Still, overall, the SOM seems to corroborate the k-means solution, as well as fit with our general, expert knowledge of Summit County.

Goal 2: Our second goal for testing $C2$ was to see what sort of causal model emerged from our cluster analysis. To do so, we used the results found in Figure 3. One of the major strengths of the SOM is that it can project all 17 variables onto the u-matrix, showing how the distribution of each variable helped to cluster and place the 20 communities in Summit County relative to one another. Figure 3 is the visual product of the SOM's variable placement. Each of the 17 small maps in Figure 3 represents the distribution of each variable on the u-matrix. For example, the map for Household Income shows that the highest household incomes are in the top left-hand corner and, moving along in clockwise fashion, the lowest household incomes are in the bottom left-hand corner. The location of the lowest household incomes also happens to be the place on the u-matrix where the highest rates of poverty, public assistance, years of life lost and so forth are located.

With the information from Figure 3 in hand, we were able to generate thick causal descriptions for each of the clusters in Figure 2. Consider, for example, Clus-

TABLE 5

Final K-means Cluster Solution for 20 Communities in Summit County

Variables (Unless otherwise noted, all data is from 1990—See Table 3)	Cluster						
	1	2	3	4	5	6	7
% Non-Hispanic Caucasian	97.3*	68.6	93.5	97.6	93.8	98.4	77.5
% African-American	1.7	28.0	5.6	1.0	4.7	1.0	21.2
% Overall Poverty	3.60	44.30	6.04	1.00	2.60	6.77	19.30
1990 household Income	41464	11404	36021	68083	49144	30002	21688
Job Growth (1993 to 2000)	31.87	20.80	17.36	27.70	43.10	15.83	.33
% Civilian Labor Force (16+ old)	96.17	85.90	95.22	96.60	95.70	94.73	90.82
% Receiving Public Assistance	2.8	25.8	4.3	1.4	2.6	5.6	13.8
% No High School Degree (25yrs+)	15.3	41.5	16.8	2.7	11.1	22.1	29.4
% of households mortgage/rent is <30% of income	16.0	43.4	17.6	15.8	19.0	18.1	27.4
% Unemployed	3.8	14.1	4.8	3.4	4.3	5.3	9.2
% No 1st Trimester Care 1995-98	5.63	24.60	7.54	1.20	4.80	8.90	14.78
Teen Pregnancies per 1000 births (1995-1998)	5.80	66.00	12.54	1.30	3.50	12.33	47.72
% children immunized by 2yrs of age	74.1	40.0	76.5	86.1	72.9	78.1	60.7
% No Health Care Coverage	4.20	25.30	6.34	1.20	3.70	8.40	14.52
Child Abuse/Neglect Rate per 1000	10.8	98.3	19.3	4.0	6.8	16.2	60.5
Elder Abuse/Neglect Rate per 1000	4.1	53.8	4.9	2.1	4.8	9.1	9.3
Years Lost per Death 1998	13.83	16.40	13.96	10.50	10.60	14.40	15.18

1. (*) The values listed in the columns for all 7 clusters represent the average value/measurement that the communities in that cluster scored for each variable listed in Column 1. In cluster analysis, these averages are called the cluster's centroids. **2. Community Membership for each of the 7 Clusters is as follows: Cluster 1**: Stow/ Silverlake, Northfield/Macedonia/Sagamore, and Richfield/Peninsula; **Cluster 2**: Central Akron; **Cluster 3**: Twinsburg, Northwest Akron, Munroe Falls/Tallmadge, Norton and Franklin; **Cluster 4**: Hudson; **Cluster 5**: Copley/ Bath/Fairlawn; **Cluster 6**: Springfield, Coventry/Green and Cuyahoga Falls; **Cluster 7**: North, West, Southwest, South and Southeast Akron and Barberton City.

ter 2, which contained the poorest community in Summit County, Central Akron. Moving from left to right across Figure 3, one sees that, in the 1990s this community (which the SOM locates in the bottom left-hand corner of the u-matrix) had the lowest percentage of whites and the highest percentage of African-Americans; the highest poverty rate; the lowest household income; one of the lowest job growth rates; the lowest workforce percentages; the highest rate of public assistance; the worst graduation record; the highest percentage of unemployment and some of the worst health outcomes indicators, including the worst mortality rate. Finally, it had the greatest number of households where the mortgage/rent was greater than 30 percent of income. In fact, when combined with the k-means information in Table 5 , one begins to develop

FIGURE 2:
Final SOM Solution for 20 Communities in Summit County

Figure 2a

(Topographical Solution for Case Study. Index for 2a on right is unstandardized distance, moving from low values (dark) to high values (light) The lighter the polygons the greater their conceptual distance is from one another.)

Figure 2b

(Cluster Solution for Case Study. The numbers in Figure 2b represent the k-means cluster solution to which each community belongs. Figure 2b is best read in clockwise fashion, moving from the most affluent and healthiest communities in the top left, to the least healthiest communities in the lower left.
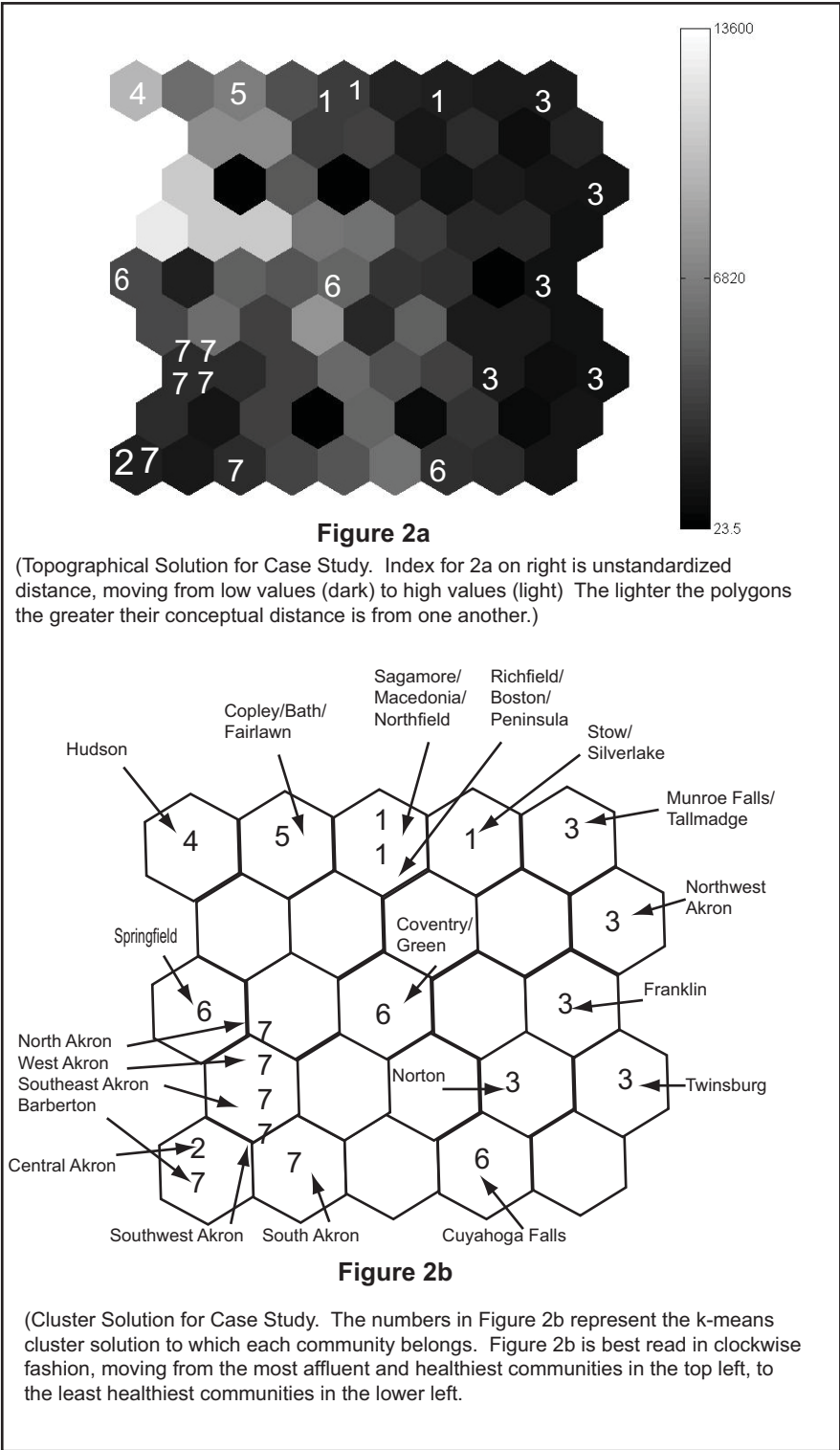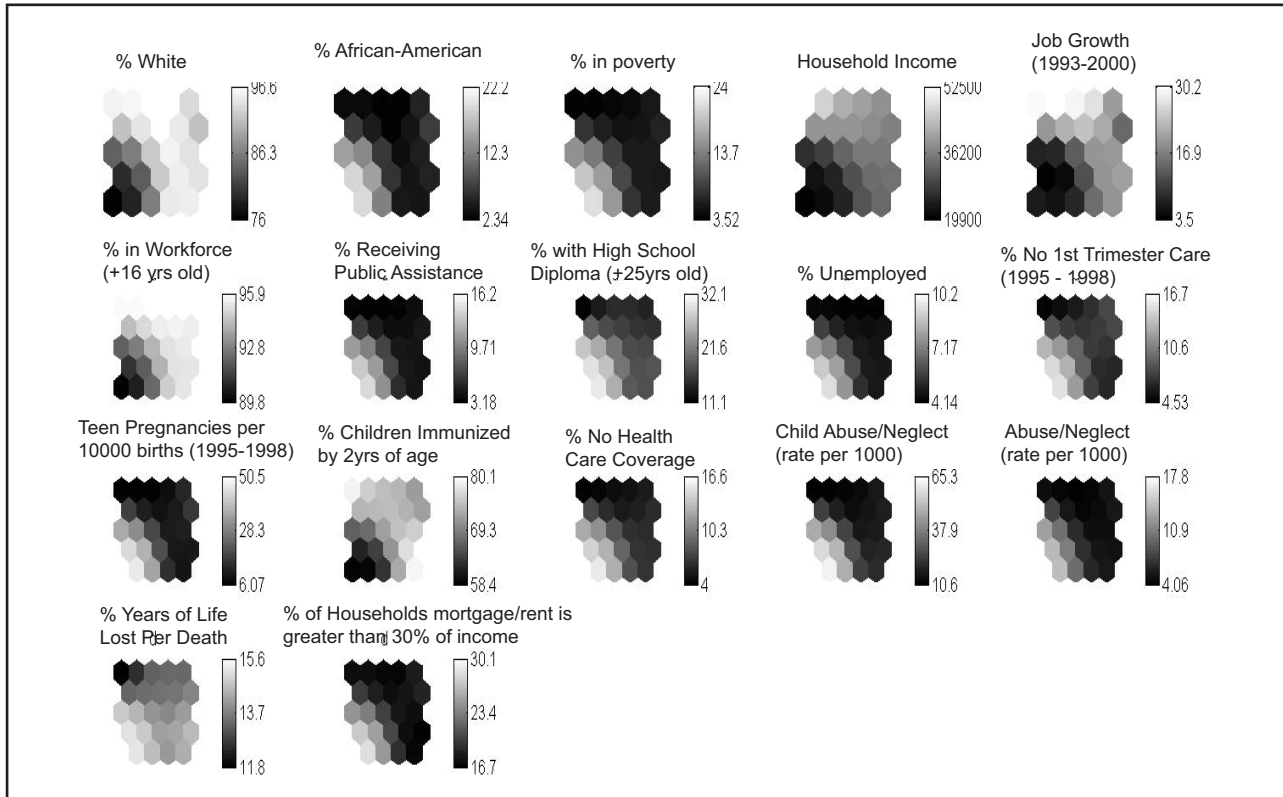
FIGURE 3:
SOM Maps of the 17 Variables Used to Cluster the 20 Communities in Case Study

NOTE: For each factor listed in Table 3, the SOM has generated a map, showing how that variable is distrubuted according to its impact on the U-Matrix in Figure 2. Like the U-matrix, the darker the polygon, the lower the value; the ligher the polygon, the higher the value. For example, the highest household income values are in the upper left of the **Household Income** map--which is where the SOM placed the most affluent communities on the U-matrix in Figure 2.

a rather sophisticated narrative of Central Akron: its median household income is roughly eleven thousand dollars; its job growth between 1993 and 2000 was 20 percent; and yet, only 86 percent of its labor force was working; and a quarter of the population was on some type of public assistance and did not have health insurance. Furthermore, the average person living in Central Akron lost 16.4 years of life, compared to Hudson, the most affluent community, which (at years of life lost = 10.5), lived an average of 6 more years. Other health outcomes were equally severe. In comparison to Hudson, which has a teen pregnancy rate of 1.3 per 1000; the rate for Akron is 66 per 1000. Less than 40 percent of children in Central Akron have received their age-appropriate immunizations by age 2; and, in comparison to the child and elder abuse rates in Hudson, which are almost nonexistent; the respective rates in Central Akron are 98.3 and 53.8 per 1000.

What is even more analytically interesting, however, is when the profile of Central Akron is compared to the communities in Cluster 7. While Central Akron has

some of the worst compositional, contextual and health outcomes, Southwest Akron, for example, despite being included in Cluster 7, has a slightly higher mortality rate—with years of life lost per death at 17.4, compared to Central Akron's 16.4. And yet, Southwest Akron was placed in Cluster 7 because, overall, it is doing better in terms of contextual and compositional factors. For example, Southwest Akron's household income level is roughly 7,500 dollars higher than Central Akron. Why, then, does Southwest Akron have a higher mortality rate? The answer is not found in any one variable. Instead, as we will explain in our test of $C4$, it is found by looking at the configuration of this community as a whole, over time. Between 1990 and 2000, the socioeconomic health and wellbeing of Southwest Akron spiraled downward—despite job and household income growth. In other words, Southwest Akron is in systems failure.

Of course, if the current article was only about $C2$ or creating narratives about the configurations of these 20 communities, we could go on to construct an increas-

ingly detailed causal model of Summit County, comparing and contrasting configurations; and we could go on to collect more detailed qualitative information to develop our narratives even further. Suffice to say for now that our results validate the main points of $C2$. Furthermore, in terms of a negative test, our results suggest that (a) $C2$ is a theoretically valuable improvement over conventional theory and method, which would struggle to arrive at such a case-based, qualitatively-subtle, variable-complex portrait; (b) $C2$ is not a repetition of what is already known; and (c) because the insights of $C2$ are narrative in form, they are easy to understand.

3.3 3. Places are Nodes within Networks

The main point of $C3$ is that the health and wellbeing of places is, in part, a function of the larger socio-spatial networks in which they are situated. As Cummins et al [19] state, "places may be more usefully viewed as nodes in networks than as discrete and autonomous bounded spatial units" (p. 1827).

As we will discuss later, studying places as "nodes in networks" links $C3$ to Characteristic 7 ($C7$): the idea that communities are open-ended with fuzzy boundaries. However, while the main focus of $C7$ is exploring "places as nodes in local, regional and transnational flows of information and other resources" (Cummins et al., 2007, p. 1832); the main focus of $C3$ is exploring "the position of places relative to each other" ([19], p. 1832). As Cummins et al. [19] state: "Studies often ignore issues of spatial autocorrelation [clustering] and assume that conditions in each locality operate on population health independently of conditions in other areas" (p. 1832). In other words, studying places as 'nodes within networks' or exploring the 'autocorrelation amongst a set of geographically proximate communities' is really about mapping and researching communities based on their relative geographical and socioeconomic position to one another.

Like $C1$, $C3$ has received empirical attention. In fact, two of the articles in Dunn and Cummins's special edition address this topic: (1) Cox, Boyle, Davey, Feng and Morris [18] and (2) Sridharan, Tunstall, Lawder and Mitchell [57]. In terms of sprawl, $C3$ hold much promise because, as we explained in our introduction, sprawl is, by definition, a network phenomenon that takes place within, between and across an evolving set of communities.

To test $C3$, we used Pajek (see methods section) to explore how communities can be conceptualized as nodes in networks and to see what additional information this conceptualization provided. To do so, we constructed a network map of Summit County and its 20 communities for 1990, using the results we obtained from our cluster analysis in $C2$. (As a side note, our test of $C3$ using 2000 data proved to be equally valid.)

We used the results from our cluster analysis for four reasons. First, if the purpose of our network analysis is to examine the configurations of communities relative to one another, we need (as shown in Table 5) the results from our cluster analysis. Second, the u-matrix from our cluster analysis provided us a spatial (albeit non-network) representation of our 20 communities relative to one another, based on their different configurations. Third, our k-means cluster analysis provided within-cluster distance measures for each of our seven clusters. These within-cluster distances are useful because they provide a weighted, spatial representation of how the cases for each cluster auto-correlate based on their position relative to each cluster's center. Finally, our k-means cluster analysis provided us with the un-standardized Euclidean distances between our seven cluster centers. The shortest distance between two points is a straight line. In cluster analysis, a Euclidean distance measure, albeit unstandardized, tells us the shortest conceptual distance between the seven cluster centers in Summit County. While, to our knowledge, nobody has used distance measures to construct a network map of communities, such data are ripe for network analysis as they are weighted, spatial representations of the relationships amongst a set of cluster centers—which are, in turn, spatial representations of the different complex configurations that exist amongst similar communities.

Figure 4 is a network representation of our cluster analysis data. The network is made up of our seven cluster centers, labeled 1 through 7. Around each cluster are the communities associated with it. Like Figure 2, the greater the distance between cluster centers, the less alike these clusters are; and, the greater the distance a community is from its cluster center, the less similar its configuration is to the other communities in its cluster.

Looking at Figure 4, the spatial positioning of the seven major clusters and their respective communities provides useful information. First, it supports Sridharan, Tunstall, Lawder and Mitchell's [57] idea that spatial auto-correlation is important. The communities in five of the seven clusters are tightly grouped around their centers, suggesting that the communities of Summit County are neither isolated nor alone in the struggles or successes with which they are dealing. For example, Pajek positioned the seven poorest communities (clusters 2 and 7) next to each other. In terms of spatial auto-correlation, all seven communities are geographically proximate to and socioeconomically in-

Figure 4:

Network Map of the Seven Clusters in Summit County and their Respective Communities
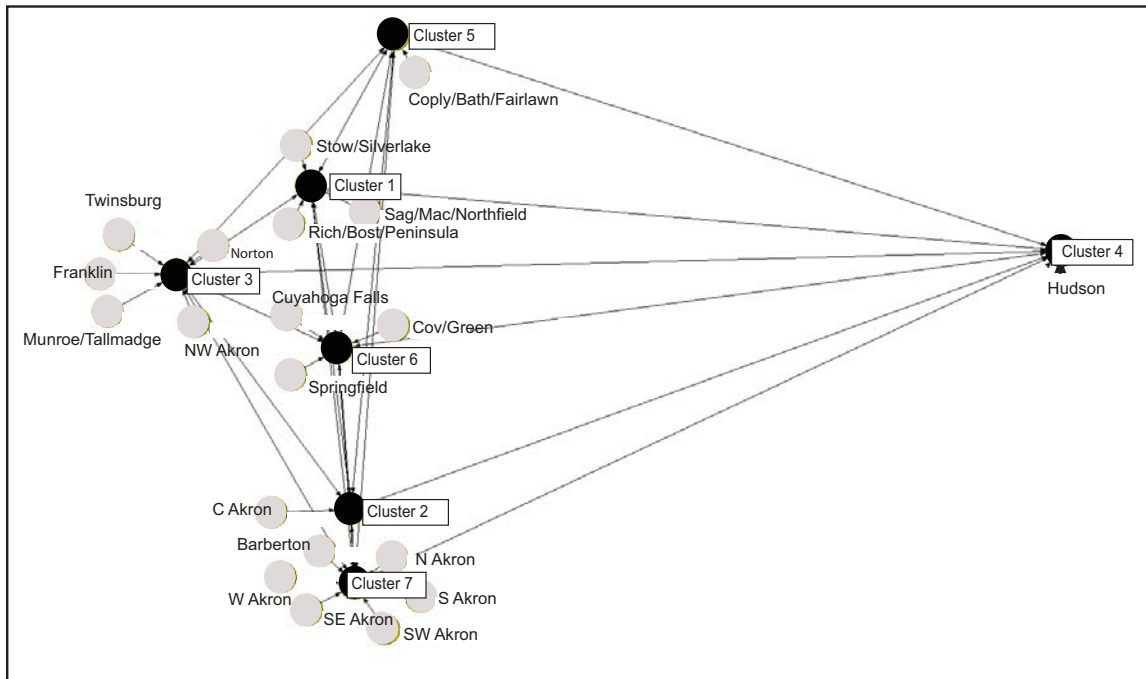
NOTE: Distances between clusters are based on Euclidian distances arrived at through k-means analysis. Distances within clusters for each community are based on within-cluster measures. All measures are non-standardized.

terdependent with each other, making up the two inter-connected urban centers in Summit County (Akron and Baberton)—see Map 1. The poverty of all seven communities is also interlinked, as Akron and Barberton struggle with issues of out-migration, the collapse of their industrial-based economies, the growing poverty of their citizens, and the failures of their public institutions, particularly education, to assuage these challenges.

The spatial position of Cluster 4 (Hudson) further corroborates $C3$, as it shows how, in terms of sprawl, the spatial wellbeing of some clusters depends upon the deprivation of others. In fact, as will be discussed in our tests of Characteristics 4, 7 and 8, the significant health and wellbeing of Hudson is linked, in part, to the economic struggles of Akron and Barberton and their surrounding communities, along with county-level residential mobility issues and suburban sprawl. Figure 4 also corroborates the findings of $C2$, particularly the spatial arrangement of the communities shown in Figure 3, which was generated by the SOM. As such, Figure 4 does a good job visually demonstrating just how qualitatively different Cluster 4 is from the rest of the communities. Finally, the position of Clusters 6, 3, 1 and 5 respectively is in accordance with their increasing health and wellbeing.

3.4 4. Places Are Dynamic and Evolving

The main point of C4 is that places are best studied as dynamic and evolving. As Cummins et al [19] state, "[I]t may be just as important for contextual studies to begin to understand not just the life course of individuals, but also the social and economic trajectories of the places which they inhabit" (p. 1832). Or, as Gatrell [28] states, "Complex systems have a history; their past is 'co-responsible' for their present behavior" ([28], p. 2662). However, while the basic point of C4 is clear enough, the methodological challenge is figuring out how exactly to model these ideas; particularly given the somewhat vague and historical manner in which the COP literature uses these terms, in comparison to the more specific and often times precisely mathematical (albeit computational) meaning.

In complexity science, to say that a place is dynamic and evolving means, at its most basic, several key things [40]. First, it means that a place is in a constant state of motion. Second, this motion emerges out of the interactions of its multiple parts (i.e., agents, forces, etc). Third, a place evolves across time/space and does so along a certain set of trajectories—which represent the set of all 'empirically possible' states a place can take at a particular moment, partitioned from the larger set of all 'theoretically possible' states (the state space).

Fourth, while dynamic, the evolution of most places falls into a relatively stable pattern, known as its set of attractor points. Sometimes, however, a place's attractor point(s) can turn chaotic or strange—think, for example, of a sudden economic collapse or moments when political unrest in a town or city becomes chaotic, or the sudden exodus of people, as in the case of sprawl. In such instances, the future state of a place is difficult to predict. But, generally speaking, the dynamics of most places, while stochastic, do evolve along a reasonably stable set of trajectories (attractor points).

The goal of our test was to examine the evolution of the 20 communities in Summit County by modeling their state space in 1990 in order to locate the dominant attractor points in the system, and then to compare these results to a second point in time, 2000, to see how this network had evolved and, more specifically, if any of the trajectories had changed; and, if so, how. We accomplished this goal as follows.

First, the dataset for the cluster analysis needed to be treated as if it were S. In our study, for example, we treated all 20 communities as if they were part of the complex system called Summit County. This makes the search for attractor points (trajectories) case-based.

Second, k-means and the SOM needed to be employed to identify, map and analyze the most common vector configurations in S for a particular moment in time/space.

Third, the obtained cluster centers (centroids/neurons) needed to be treated as S's attractor points. In our study, for example, the seven clusters discussed in our results section became the attractor points in our system (Summit County), around which the 20 communities in our study grouped. The un-standardized Euclidian distance measures provided in the proximity matrix for S need to be mapped to visually identify S's attractor points and the cases clustering around them. In our study, for example, Figure 4 becomes our map of Summit County and its attractor points for 1990. We generated this map by entering our unstandardized weights into Pajek, a network analysis software package. The map was generated using the popular network visualization algorithm, Kamada-Kawai [36].

Fourth, the within-cluster distance measures for the cases needed to be used to map the relative distance of each case to the particular cluster/trajectory/attractor point to which it belongs. This is useful because the cases clustering around the system's attractor points become data for creating a thick description of the various trajectories toward which the system is drawn. In our study, for example, the communities clustering around each solution were used to construct thick descriptions of the different directions Summit County is

heading. Here is a very brief description of what Figure 4 suggests: Summit County seems to have evolved into three, main trajectories: there is a trajectory toward affluence and health, represented by Cluster 4; a trajectory toward poverty and poor health, represented by clusters 2 and 7; and a sort of middle ground trajectory, perhaps representing the major settling point for this County, which revolves around average to above average health and wellbeing, represented by clusters 1, 3, 5 and 6.

Although not shown in Figure 4, to further test our notion that clusters 1, 3, 5 and 6 constitute the major settling point for this County, we re-ran our cluster analysis, entering Summit County as a 21st community, using for its vector configuration the County-level averages for all 17 variables. Our cluster analysis grouped Summit County with Cluster 6.

Finally, if such information is available, the above five steps can be repeated for additional points in time. These additional points in time can be then compared to the first point in time to see how the attractor points in the system might change.

For example, for our study we re-ran the above analyses for 2000, to explore how the 7 trajectories in Summit County changed over a ten year period. We used the same entry order as shown in Figure 4, using 2000 data for all compositional and contextual factors. Our health outcomes, however, did not change, as these were aggregated across varying periods of time between 1990 and 2000. For example, Years of Life Lost per Death came from data aggregated between 1990 and 1998. Running our k-means we also sought the same 7-cluster solution found in 1990 to see if it continued to prove useful. Table 6 is a quick summary of our results. In terms of reading Table 6, the second column shows cluster membership for each of our 20 communities in 1990; the third column shows their 2000 membership. Also, going from the top of the table to the bottom, we ordered the clusters from the most affluent to the poorest.

The first thing that stands out in Table 6 is that the seven major trajectories in Summit County continued to exist in 2000. However, there were some interesting shifts. For example, the poorest trajectory, represented by Cluster 2, gained a community, Southwest Akron. On the opposite side, the second most affluent trajectory, represented by Cluster 5, also gained a community. These two shifts suggest that the richest and poorest trajectories not only gained in strength but also, between 1990 and 2000, a widening socioeconomic gap emerged between the poorest and richest communities. This gap was further corroborated when we re-ran our 2000 k-means with Summit County included. The County dropped from Cluster 6 to Cluster 7, which con-

TABLE 6
Change in Final Cluster Solutions for 20 Communities in Summit
County, 1990 to 2000

COMMUNITY	YEAR	
	1990 Cluster Membership	2000 Cluster Membership
(Affluent Cluster) Hudson	4	4
(Affluent Cluster) Copley/Bath/Fairlawn	5	5
(Middle Class Cluster) Stow/Silverlake	1	1
Northfield/Macedonia/Sagamore	1	1
Richfield/Peninsula	1	**5***
Twinsburg	3	**1***
Northwest Akron	3	3
Munroe Falls/Tallmadge	3	3
Norton	3	6
Franklin	3	3
Springfield	6	6
Coventry/Green	6	**3***
Cuyahoga Falls	6	6
(Poor Cluster) North Akron	7	7
West Akron	7	7
South Akron	7	7
Southwest Akron	7	**2***
Southeast Akron	7	7
Barberton City	7	7
(Poorest Cluster) Central Akron	2	2

1. (*) The values listed in the columns for all 7 clusters represent the average
value/measurement

tains most of the poor communities and, after Cluster 2, the worst health outcomes.

There is a lot more to be explored in our analysis of the evolution of the network of communities in Summit County. However, given the need to report on our second test, we must stop here-content with the fact that the test accomplished its goal: to determine if the evolution of a network of communities could be effectively modeled with the SACS Toolkit and if such an analysis could provide theoretically novel insights into the evolution of this complex system.

3.5 5. Places are Nonlinear.

Scholars involved in the study of place are very clear about how they define and intend the characteristics of nonlinearity. As Gatrell [28] and others explain (e.g., [22], [23], [26]), in terms of a complex system like place, nonlinearity addresses the empirical fact that, more of-

ten than not, small or large changes in some aspect of a place (e.g., its health system, educational system, etc), particularly in the form of health interventions (e.g., new outpatient program, new educational accountability measures, etc) do not regularly lead to their expected, linearly related outcomes (e.g., a 25 percent increase in prenatal care or graduation rates, etc).

Our test of $C5$ attempted to examine if it is realistic to assume that, given the complexities of sprawl, and its unequal impact on community-level health, that health interventions by a County to improve its communities are, indeed, nonlinear in their outcomes. For our test, we focused on a report released by the Summit County Social Services Advisory Board (SSAB), titled Summit 2010 Priority Indicators Progress Report, 2009 (www.healthysummit.org/QOL/QOL-About.cfm).

In 2003, the SSAB began a program it called Healthy Summit 2010, a community-level version of the federal government's Healthy People 2010. The purpose of Summit 2010 was to establish a set of 20 goals for improving the economic, institutional, physical and behavioral health and wellbeing of Summit County. In turn, as shown in Table 7, these 20 goals were turned into priority indicators that the SSAB could use to measure the County's progress.

In terms of testing the concept of nonlinearity, two aspects of the Healthy Summit 2010 project need to be addressed: effort and outcomes. In terms of effort, it is clear from reading the Healthy Summit 2010 Quality of Life Project website that the work being done in Summit County to improve the health and wellbeing of its citizens, particularly those in need, is ambitious both in scope and in effort. This ambitious, concerted effort includes the SSAB; its numerous committees; directors and researchers; an extensive network of health and social service systems; and a long list of health care providers, politicians, community leaders and activists. In fact, as stated on the front page of the website: "The Quality of Life Project, under the guidance of the Social Services Advisory Board, will lead Summit County to new, unprecedented success in the age-old battle to improve health, expand economic opportunity, and reduce poverty and its ill effects."

The challenge, however, is that, between the years of 2003 and 2009 there was no statistically significant, linear relationship between the concerted efforts of the Healthy Summit Quality of Life Project and their 20 health outcomes. In fact, looking at Table 7, the data show that improvements across the Healthy Summit 2010 twenty key indicators were, at best, uneven. For example, while a few indicators, such as Education (Indicator 4) and Teen Birth Rate (Indicator 13) showed some level of improvement, other indicators, such as

General Poverty (Indicator 1) and Unemployment (Indicator 3) got worse; and other indicators, such as Years of Potential Life Lost (Indicator 20) remained the same.

As stated previously, the nonlinear relationship between effort and outcome is a major challenge for community health providers (e.g., [38]). It is usually unclear, for example, what types of concerted efforts will yield the measureable outcomes desired. For example, while research on chaos theory (e.g., [31]) and tipping points (e.g., [30]) suggests that small changes in initial conditions can lead to sudden and significant change in outcomes; research on poverty traps (e.g., [7]) and nonlinearity in complex human organizations [48] suggests that significant efforts often lead to little difference. Related, it is not always clear how best to measure change. For example, will a community's efforts to address poverty yield five years of no results, only to suddenly produce significant change by year six? While we cannot address these questions in the current study, our quick test of Summit County suggest that, in complex systems such as communities the relationship between effort and outcome can be rather nonlinear.

3.6 6. Places Are Subjective and Historical

Like $C1$, $C6$ has a positive and critical form. In its positive form, its point is that subjectivity, personal experience and history play an important role in the dynamics of places. This part of $C6$ is self-evident and does not need testing. The critical form of $C6$, however, does need to be addressed. It basically argues that, while scholars involved in the study of place know that history and subjectivity play an important role in health, they ignore this type of research, opting instead for quantitative analysis. As Cummins et al [19] explain: While a call has been made to explore health as the "lived" and "embodied" experiences of people interacting with their settings, these ideas "remain poorly integrated into empirical research" (p. 1829). To address this lack of integration, researchers need to incorporate into their work "information about settings that are drawn from reported views of residents, as well as from independently measured indicators of local conditions" ([19], p. 1830).

While readers may not know, complexity scientists almost unanimously share the same "anti-qualitative" bias as the study of place. In fact, following Smith and Jenks[55] and others (e.g., [13], [15]), there is almost no qualitative or historical research done in complexity science. For example, as Gatrell [28] states, "[T]he human voice seems to be missing from much of the complexity theory. The qualitative is there, but in the form of qualitative structures and patterns, not in the nature of

Table 7: Health Outcome Indicators for Healthy Summit 2010

Priority Indicator	Description	Indicator currently better, same, or worse than in 2003 Environmental Scan
Indicator 1	Increase the proportion of people living above the official poverty line	**Worse**
Indicator 2	Increase the proportion of African-Americans living above the poverty line	**Worse**
Indicator 3	Reduce unemployment	**Worse**
Indicator 4	Increase the proportion of people aged 25 and over who have received a high school diploma	**Better**
Indicator 5	Increase housing affordability, raising the proportion of households spending under 30 percent of their incomes on housing	**Worse**
Indicator 6	Reduce the proportion of households receiving Temporary Assistance for Needy Families (TANF)	**Better ****
Indicator 7	Reduce the incidence of domestic violence-related crime	**Worse**
Indicator 8	Reduce the rate of violent crime	**Comparable data currently unavailable**
Indicator 9	Increase the proportion of African-American children under age 5 living above the official poverty line	**Same**
Indicator 10	Increase the proportion of children receiving their immunizations by their second birthdays	**Worse**
Indicator 11	Reduce the incidence of child abuse and neglect	**Better**
Indicator 12	Increase secondary school attendance	**Better**
Indicator 13	Increase the proportion of African-American children age 18 or less living above the federal poverty level	**Same**
Indicator 14	Reduce the rate of births to teens, focusing on higher rates among African- American youth	**Better**
Indicator 15	Increase the proportion of African-American older adults (age 65+) living above poverty	**Same**
Indicator 16	Increase self-sufficiency of seniors living alone	**Same**
Indicator 17	Reduce the incidence of elder abuse and neglect	**Same**
Indicator 18	Increase the proportion of individuals with health insurance	**Worse**
Indicator 19	Increase the proportion of pregnant women receiving first trimester prenatal care	**Better**
Indicator 20	Reduce the rate of Years of Potential Life Lost from All Causes	**Same**

Better - The indicator improved relative to the 2003 Environmental Scan; *Same* - The indicator did not show any appreciable change from the 2003 Environmental Scan; *Worse* - The indicator declined relative to the 2003 Environmental Scan; *Unknown* - Because of missing data or overlapping confidence intervals, it is unknown whether any change in the indicator occurred.

** While the goal of reducing the percentage of the population on public assistance has been met, the "improvement" seen is most likely being caused by needy families using up their current eligibility for public assistance rather than an increase in economic self-sufficiency.

the embodied actor" (p. 2669). And so, the inclusion of $C6$ in Table 1 constitutes another important addition to the quasi-general definitions of complex systems typically transported into the social sciences.

Given the self-evident nature of $C6$, the purpose of our test was to see if qualitative information improved our theoretical understanding of Summit County and the impact sprawl is having on it. For our test, we used two qualitative reports located on the Healthy Summit 2010 Quality of Life Project website Visit the website at: www.healthysummit.org/QOL/qol-about.cfm).

First, there is the Neighborhood Project Summary Report. This report summarizes a series of focus groups conducted with members living in three targeted neighborhoods in Summit County: Barberton, Buchtel and Lakemore. Second, there is the Report on Key Informant and Community In-person Interviews. This report was based on N=230 interviews: 190 interviews with key informants from the public, nonprofit and private sectors in Summit County; and 40 interviews with county residents—two from each of the 20 communities in Summit County. We used these reports because they are qualitative in nature, giving attention to issues of voice and subjective experience; and because these reports were used to develop the Healthy Summit 2010 Quality of Life Project, including its strategic interventions for improving health.

To test $C6$ we read through these two reports to determine if an appreciation of history, subjective experience and voice added important information to our understanding of Summit County and its health. Here is a quick example of some of our insights.

The Neighborhood Project Summary Report (NPSR) engaged local citizens in three neighborhoods to establish a set of goals for addressing the major health care problems they face. This data revealed for example that, in terms of addressing health disparities in Barberton (one of the poorest communities), "The primary concerns of the residents, elected officials, and professionals uniformly 'centered around' employment. Educational concerns, family problems, and crime were all seen as directly related to employment difficulties. The loss of jobs in the city, education not matching employment opportunities, and difficulties with transportation to employment and training opportunities were cited as areas needing intervention" (NPSR, p. 10). Note that these issues are directly tied to sprawl: as jobs, affluence and resources leave these inner-city communities, we see that work, education and transportation become major issues.

In turn, the Report on Key Informant and Community In-person Interviews collected first-person views on a host of political, economic and cultural issues in Sum-

mit County. Several findings had a lot to say about the underlying causal mechanisms related to sprawl that were responsible for the simultaneous collapse of urban communities and rise of some suburban communities, including: (a) racism; (b) the negative perception affluent communities have of public services; (c) affluent flight from Akron to the suburbs; and (d) political and economic turf battles between certain communities as affluence moved out into the suburbs. The same report also solicited opinion on the strengths and weaknesses, opportunities and threats for a host of health issues in some of the poorer communities, including (a) effectiveness of community-based organizations; (b) children and youth programs; (c) services for working adults; (d) family services and (e) elderly care. While we cannot go further in our review of these reports, suffice to say we found them rich with information about the historical, political, economic, cultural and institutional nuances of Summit County and its health and health care, allowing us to conclude that this type of information is a valid and valuable part of understanding the structure and dynamics of communities as complex systems.

3.7 7. Places are Open-ended with Fuzzy Boundaries

The main point of $C7$ is that places are not autonomous entities or closed-systems with clear geographical or sociological boundaries. Instead, they are open systems with fuzzy boundaries. Like many of the characteristics in Table 1, $C7$ is well defined. Let us explain.

In complexity science, an open system is one that interacts with its environment. Almost all biological life forms are, in principle, open systems insomuch as they survive through their adaptive interactions with the environmental mediums in which they live ([14], [46]). Places are similar. However, in the case of contemporary, globalized society, survival not only depends on a place's interactions with the physical environment, but also the larger globalized, economic, political and cultural systems in which it is situated.

The open-ended nature of places takes us back to $C3$ and the study of social networks. As a reminder, while $C3$ focused on communities as "nodes within networks," $C7$ focuses on communities "as nodes in local, regional and transnational flows of information and other resources" ([19], p. 1832). For example, Cummins et al. state, "We can conceive, for example, that trends [flows] in regional economies, national and regional environmental pollution, national or supra-national organizations and entities can all define the 'local' and other contexts in differing ways and this in turn contributes to the spatial distribution of health outcomes" (p. 1833).

In addition to being open-ended, the boundaries of communities tend to be fuzzy. Fuzziness has to do with mobility—perhaps the most important factor that engines sprawl. In a global society, the agents in a community (be they people, businesses or health care systems) often do not confine themselves to the geographical or sociological boundaries of their home communities [59]. Health care systems, for example, tend to spread out, building their centers in other communities or caring for patients within a particular region. In turn, in sprawling counties like Summit County, people often live in one community while working in another; get their groceries in one place while receiving their health care in another. The social networks of people also bleed across communities, especially when one allows for the influence of cyber-infrastructure (e.g., cell phones and online social networks). We could go on. The point is that mobility is a major issue in terms of defining the boundaries of places, so much so that it is best to: (a) view place boundaries as fuzzy and (b) define the "place-level" exposure of many individuals as multiply determined. As Cummins et al [19] state, "These issues of varying individual-level exposure to multiple contexts over time and space means that current measures of simple universally applied 'neighborhood' exposure may severely underestimate the total effect of 'context'"(p. 1830).

To test $C7$ we turned to a study recently completed by the Northeast Ohio Urban Sprawl Modeling Project (NOUSMP)—See Peterson et al [49]. The purpose of NOUSMP was to provide the general public a data-driven visualization of the projected impact that urban sprawl will have on the 15 counties of northeast Ohio by 2020, including Summit County (See p. 2-4).

Figure 5 was created from maps generated by Arc-Explorer, a public domain software program for viewing and studying geographical information systems (GIS) data. The GIS data for Figure 5 came from a variety of environmental and geospatial databases (p. 2-14), which NOUSMP researchers entered into their model. The projected spatial distribution of Summit County's population, as shown in Figure 5 Map B, was based on a simple set of assumptions about how urban sprawl would occur between 1990 and 2020, such as projected residential growth along highways (p. 2-4). Putting all of this together, Figure 5 shows the change in population density in Summit County as a function of urban sprawl in northeast Ohio (circa 1990, Figure 5, map a) and its projected impact in 2020 (Figure 5, map b).

Here is how Figure 5 helps us test $C7$. Maps A and B in Figure 5 are organized according to the major communities of Summit County. The lines moving out of Akron denote the major highways and roads in Summit County. In terms of population density, the darker the color, the denser the population. Looking at Map A, the highest rates of population density in Summit County are found in the communities surrounding the city of Akron, with almost no major density along any of the highways. Map B shows a different story. By 2020, the population density of Summit County shifts away from the cities, concentrating around most of the major highways and roads of the second-tier suburban communities of Summit County—see the arrows in Map B.
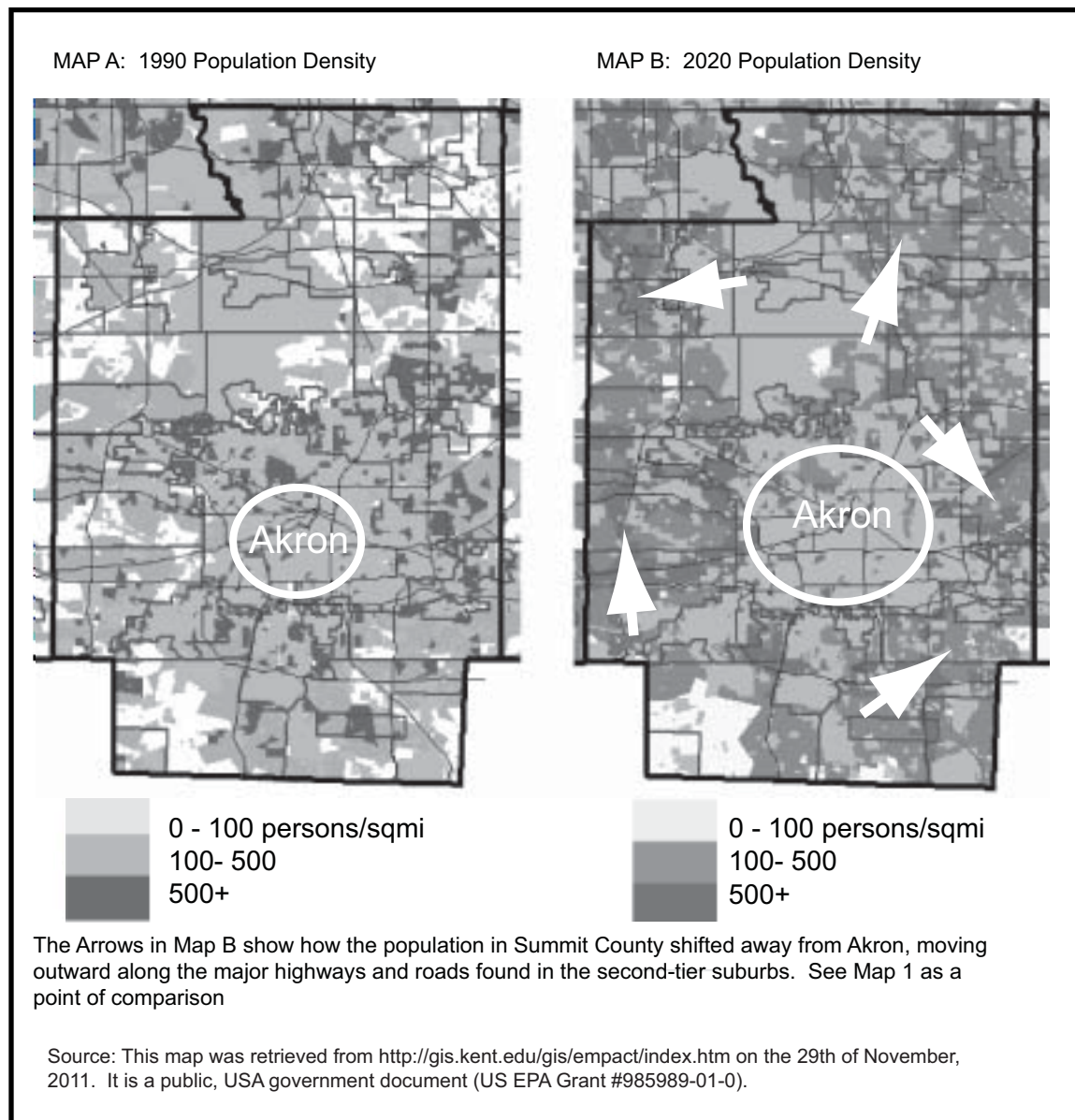
The shift in population density shown in Figure 5 is a small window into how the boundaries of the communities of Summit County have become increasingly open-ended and fuzzy. In terms of their open-ended nature, sprawl has forced the communities of Summit County (as nodes within a larger network) to adapt to the wider regional flows of commerce, people and capital, as well as the impact these sprawling flows have on the 'local' context. And, in terms of fuzziness, as people and jobs move into Summit County's suburban communities, they leave behind the poor. These insights happen to corroborate the 'first-person' narratives we discussed in $C6$, where outmigration and the movement of jobs elsewhere are seen as a major barrier to improving the economic wellbeing of poor communities in Akron. We could go on, but we need to stop. Our brief review of sprawl does suggest, however, that thinking about places as open-ended and fuzzy is valid and valuable.

3.8 8. Places Are Power-based, Conflicted Negotiations

$C8$ constitutes another addition made by community health science scholars to the quasi-generalist definition of complex systems used in complexity science. It is clear in its focus: places are imbued with power relations. Its empirical study, however, and its epistemological-theoretical frame need to be developed—particularly given the availability of different theories of power, from Marx to Foucault. Nonetheless, despite its tentative outline, $C8$ makes an important point that complexity scholars involved in the study of place need to address. Let us explain.

For a sociologist to state that places are imbued with power relations is not, at first glance, saying anything new. In fact, only the opposite sort of statement, the idea that places are not imbued with power relations, would require empirical test. Interestingly enough, what is obvious to one field of study or discipline is not necessarily obvious to another. Such is the case in complexity science. For all their 'sociological' discussions about the complexity of businesses, governments, financial markets, social networks, small group dynamics and communities, the majority of scholars in complexity science—particularly those coming from the natural

FIGURE 5
Maps of Summit County Population Density (1990 to 2020) as a Function of Urban Sprawl

MAP A: 1990 Population Density   MAP B: 2020 Population Density

0 - 100 persons/sqmi
100- 500
500+

0 - 100 persons/sqmi
100- 500
500+

The Arrows in Map B show how the population in Summit County shifted away from Akron, moving outward along the major highways and roads found in the second-tier suburbs. See Map 1 as a point of comparison

Source: This map was retrieved from http://gis.kent.edu/gis/empact/index.htm on the 29th of November, 2011. It is a public, USA government document (US EPA Grant #985989-01-0).

and artificial sciences—have not developed a vocabulary for discussing the impact that power relations (particularly inequality, oppression or exploitation) have on the structure and dynamics of complex systems. And, there is even less discussion of related power issues such as ethnicity, social class or gender. For example, as Gatrell states, "Gender too seems to be a missing strand from existing uses of CT [complexity theory]" ([28], p. 2669). The only caveats to this dominant trend are a handful of scholars working at the intersection of sociology and complexity science (see [15]) and, in terms of the current study, scholars working at the intersection of complexity science and community health. At second glance, then, $C8$ is less a reminder to social scientists and more a challenge to complexity scientists.

To test of $C8$ we focused on a conflict that expressed well how sprawl, as a complex systems problem, results in unintended health inequalities at the community level. The conflict has to do with the two major health systems in Summit County, the Summa Health

System (SUMMA) and Akron General Health System (AGHS). The conflict (which began circa 2008) centered on SUMMA's desire to build a 100-bed, physician-owned hospital near Hudson, the most affluent community in Summit County. Historically, the majority of facilities run by SUMMA and AGHS have resided within and around the city of Akron, which has been to the health advantage of Akron residents, particularly the poor. The building of a new hospital near Hudson and the affluent communities around it therefore signaled a worrisome change to the communities in Akron.

In response, on October of 2008, AGHS held a forum entitled, "Should Doctors Own Hospitals? And Why Should You Care? A Close Look at Risks and Realities" [44]. In their forum, AGHS explained that its main concern was that SUMMA's new facility, given its location and for-profit status, would pull physicians from the Akron area; physicians would also be able to "cherry pick" their patients, and those affluent patients out-migrating to or presently living in Hudson and the surrounding affluent suburban communities would stop going to Akron to receive care—all of which would negatively impact community hospitals and clinics in Akron, primarily in the form of "huge drops in volumes and revenue," "lowered bond ratings," and "staff layoffs and reductions," which would lead to poorer health outcomes [44]. SUMMA countered, arguing that no such thing would happen; instead, the new facility would increase access for residents migrating to or currently living outside of Akron, as well as stimulate job growth for Summit County. In terms of the current study, as might be expected, the political and economic leaders of the suburban communities vying for the location of SUMMA's new hospital saw the project in entirely positive terms; while the communities that perceived the project as problematic—those within clusters 2 and 7 of our research—saw it as corrosive to the health and wellbeing of Akron communities and, on the whole, to Summit County. As of 2011, SUMMA's 100-bed facility remains to be built.

While we cannot go into greater detail about the conflict between SUMMA and AGHS and their respective communities, it illustrates the main point of $C8$, as it relates to the issue of sprawl: places and their health are imbued with power relations that can have a significant impact of the health of their respective residents. We turn, now, to our final characteristic.

3.9 9. Communities are Agent-Based

Like $C2$, $C9$ draws on (imports) one of the most important characteristics (areas of study) in complexity science: the idea that complex systems emerge out of the self-organizing, adaptive interactions amongst a set of rule-following agents.

To say that places are agent-based means five things to community health scientists. First, it means that interacting agents play a major role in the self-organization, emergence, structure and dynamics of places and their health.

Second, it means that "places are produced and maintained by the activities of 'actors,' proximate or distal to a particular place, who operate individually or in concert across a wide range of geographical scales" ([19], p. 1828). This is a particularly useful point in terms of sprawl, which concerns the impact micro-level actions at a distance (such as moving to the suburbs) has on community health in another location.

Third, it means that it is impossible to think about compositional and contextual factors without considering how they are enacted through the complex interactions of the agents involved in a place. Again, another key point for studying sprawl and health.

Fourth, and more specifically, it means that "place effects on health emerge from complex interdependent processes in which individuals interact with each other and their environment and in which both individuals and environments adapt and change over time" ([2], p. 1). In terms of sprawl, this is very important because it helps us understand the role agency plays in the evolution we saw in our test of $C4$.

Finally, it means that agents "can be conceived of in a variety of ways from individuals and community organizations, firms and businesses, regional and national governments and institutions, peer-networks and families to static and dynamic regulatory structures and processes such as national tax policy and the rule of law" ([19], p. 1828).

In its critical form, the main point of $C9$ (like $C1$ and $C2$) is that multi-level analysis (regression) is an insufficient method for studying the impact that compositional and contextual factors have on health because it cannot account for the role that heterogeneity, agency and interaction play in the structure and dynamics of communities. As Auchincloss and Diez Roux state, "In general, regression approaches continue to be ill equipped to investigate the processes embedded in complex systems characterized by dynamic interactions between heterogeneous individuals and interactions between individuals and their environments with multiple feedback loops and adaptation" ([2], p. 2). As a result, the limitations of multi-level regression "have constrained the types of questions asked [by community health scientists], the answers received, and the hypotheses and theoretical explanations that are developed" ([2], p. 1). Agent-based modeling, however, act-

ing as a complementary method, can help to overcome these limitations.

To test the validity of C9, we created an agent-based model called Summit-Sim—see Method Section for detailed overview of how we built this model. The goal of Summit-Sim was to test a specific aspect of sprawl: residential migration patterns.

In addition to the report reviewed in our test of $C7$, the empirical basis for how we built Summit-Sim came from the Summit 2010 Priority Indicators Progress Report, 2009. In this report, sprawl is used to explain differences in community-level poverty rates across Summit County. To understand the link between sprawl and differences in community-level poverty, we need to review Map 2, which was included in the report.

Map 2 is read as follows. Dots represent the spatial location of jobs in Summit County and its larger region, Northeastern Ohio, which includes the City of Cleveland and a total of eight counties, including Summit. Each dot represents five jobs. Shaded areas on the map correspond to changes in residential density across the communities in Summit County and Northeastern Ohio. To compute residential density change, researchers looked at where people were living in 2000 compared to where they were living in 1995, arriving at a percentage of the 2000 population in a community that is new. Looking at Map 2, the northeastern suburbs of Summit County appear to be the major "hot spots" for residential migration within the Northeastern Ohio region.

Map 2 corroborates or helps explain several of the findings we discussed in several of our previous tests. First, while Map 2 provides a different picture than Figure 5 ($C7$), it supports the simulated and projected findings of the Northeast Ohio Urban Sprawl Modeling Project: Summit County is a microcosm of the larger region, as residents are moving out of the urban centers of Northeastern Ohio into the suburbs. Second, Map 2 helps explain the findings at the end of $C4$ (as a reminder, the communities in clusters 4 and 5 pulled further away from the rest of the County in 2000; and Cluster 5 gained Richfield, Boston Mills and Peninsula) by suggesting that this spatial inequality may be due to affluent residents moving into these suburban areas. Third, Map 2 supports the argument made by Akron General Health System in $C8$ that unplanned growth—specifically the movement of for-profit hospitals and health care into the northeastern suburbs of Summit County—may negatively impact the health of residents living in Akron, as money and services move out of the urban areas and into the suburbs.

Together, then, Map 2 and the results of $C4$, $C7$, and $C8$ can be used to understand how residential (mo-

bility) migration patterns in Northeastern Ohio impact differences (inequalities) in the health and wellbeing of the communities in Summit County; in particular, they can possibly explain differences between the health of the most affluent, suburban communities (clusters 4 and 5) and the poverty and declining health of the poorest urban communities (clusters 2 and 7).

Stated in formal theoretical terms, the above results on sprawl suggest that the agent-based (mobility) migration behaviors of a set of heterogeneous agents (residents) living in Northeastern Ohio changed the residential composition of Summit County between 1995 and 2000. This change occurred, in large measure, as more affluent agents migrated to more affluent suburban communities, leaving behind their less affluent neighbors. In turn, these residential migration patterns resulted in a spatial segregation of health: poor agents remained confined within or found themselves migrating to poor neighborhoods with poor health outcomes (clusters 2 and 7), while more affluent agents migrated to suburban areas with high health outcomes (clusters 4 and 5). Such a trend has the possibility of being reversed if Summit County addressed the challenges of sprawl, as outlined in $C7$, and the health consequences of this sprawl, as suggested in $C8$.
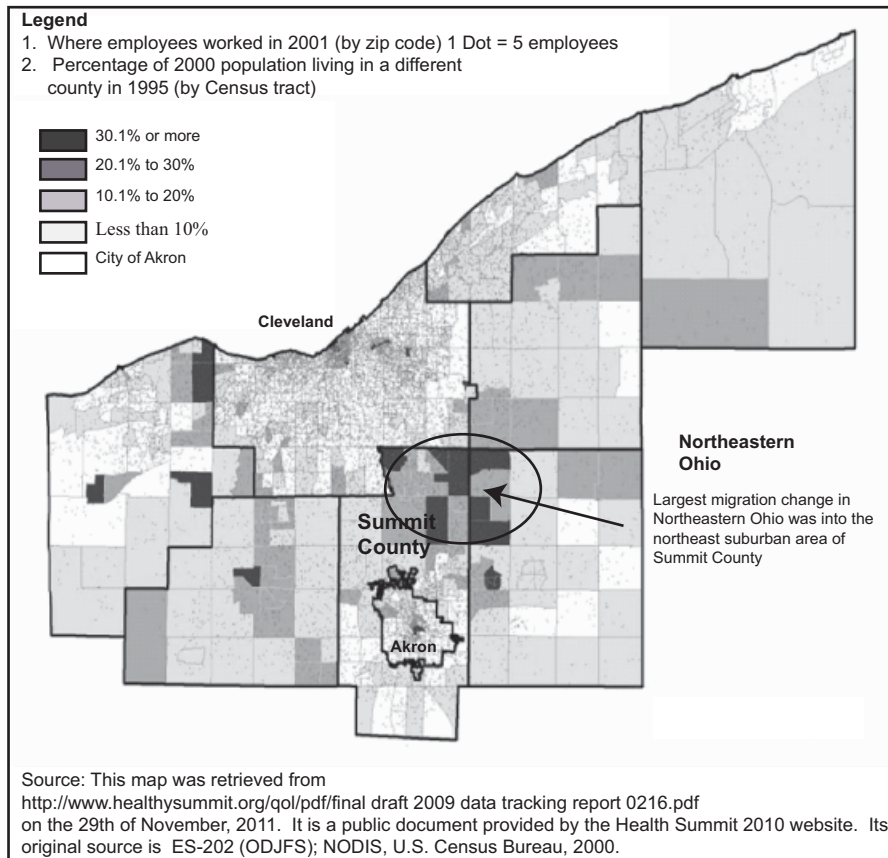
3.9.1 Enter Agent-Based Modeling

While the above theoretical statement is empirically rigorous, it would be very useful to somehow experimentally test its fundamental assumptions. But, how does one do that? Enter agent-based modeling ([27], [29]). Agent-based modeling is useful to community health science because it can experimentally verify empirically-derived theories on the relationship between context, composition and health [2]. More specifically, it is useful because it is good at modeling how the macro-level patterns of communities (e.g., residential patterns and community-level health) emerge out of the nonlinear, dynamic, micro-level behaviors of their interacting and intersecting heterogeneous agents (e.g., sprawl and residential migration behaviors). Having made this point, we turn to our model.

To test $C9$, we explored two questions. First, do the sprawling migration patterns of the heterogeneous agents living in Summit-Sim result in the clustering and spatial segregation (distribution) of affluence we see in Summit County? As shown in Figure 7, the answer is yes.

Figure 7 is a snapshot of Summit-Sim with a preference rating of 3 for all agents. This rating means that, for each iteration of Summit-Sim, rich agents sought to live in a neighborhood with three or more rich agents;

MAP 2
Change in Residential Mobility Between 1995 and 2001

Legend
1. Where employees worked in 2001 (by zip code) 1 Dot = 5 employees
2.   Percentage of 2000 population living in a different
     county in 1995 (by Census tract)

- 30.1% or more
- 20.1% to 30%
- 10.1% to 20%
- Less than 10%
- City of Akron

Cleveland

**Northeastern
Ohio**

Largest migration change in
Northeastern Ohio was into the
northeast suburban area of
Summit County

**Summit
County**

Akron

Source: This map was retrieved from
http://www.healthysummit.org/qol/pdf/final draft 2009 data tracking report 0216.pdf
on the 29th of November, 2011. It is a public document provided by the Health Summit 2010 website. Its
original source is ES-202 (ODJFS); NODIS, U.S. Census Bureau, 2000.

middle-class agents also sought to live in a neighborhood with three rich agents; if they could not migrate to such a neighborhood, they sought to live near other middle-class agents; if they found themselves in a neighborhood with four or more other middle-class agents, they stayed; finally, poor agents sought to live in neighborhoods with three or more middle-class agents if they could; if they could not, they stayed where they were.

Of the various preference ratings available for our model, we chose 3 because it is a rather modest preference. What made Schelling's model of segregation so powerful is that macro-level patterns of significant segregation resulted from very mild preference ratings. Sprawl seems to follow a similar pattern. Mild levels of neighborhood preference should produce significant spatial clustering and segregation.

A visual inspection of Figure 7—which, given the constraints of time and space, will suffice for our analysis of Summit-Sim—shows that, as expected, a preference rating of 3 leads to significant spatial clustering and to even more extreme patterns of segregation than that found in our empirical analysis of Summit County. There are very tight clusters of rich agents (see Cluster

A), surrounded by a few middle-class agents; there are large, loose clusters of middle-class agents spread out in the same basic area, moving from the top-right corner of Summit-Sim to the bottom lower-left corner. Finally, there are tight clusters of poor agents (see Cluster B), some of which are very large.

The second question we explored was: If the sprawling migration behaviors of our heterogeneous agents leads to spatial clustering, does this segregation of affluence result in community-level health inequalities, as seen in Summit County? As shown in Figure 7, the answer is yes. From the start of the model to its completion, the context-dependent unhealthiness of poor agents never dropped below roughly 50 percent. Meanwhile, the rich agents had near perfect health. These healthiness ratings fit with our analysis of Summit County; in particular, our comparison of the poorest clusters with the more affluent.

While our abbreviated analysis of Summit-Sim leaves numerous issues unexplored—for example, how do different preference ratings impact spatial segregation or health?—it is adequate to support the point of $C9$: the health and wellbeing of communities is sufficiently

FIGURE 7:
Snapshot of SummitSim with a Preference Rating of 3 for all Agents

NOTE: Rich Agents = Squares; Middle Class Agents = Stars; and Poor Agents = Triangles. **Cluster A** identifies one of the dense clusters of rich agents. **Cluster B** identifies one of the dense clusters of poor agents; which complexity scientists would call a poverty trap.

agent-based, such that studies of composition and context should include some form of agent-based analysis. We turn, now, to the conclusion of our study.

4 CONCLUSION

Considering all nine tests together, here are our conclusions regarding sprawl and health. First, in terms of the individual tests, $C1$ demonstrated the limitations of linear modeling for studying sprawl and health; $C2$, in contrast, demonstrated strong support for case-based modeling; $C3$ and $C7$ illustrated the utility of network analysis for studying sprawl as a complex network of places; $C4$ showed strong support for exploring how a sprawling network of places evolves across time/space; $C5$ shows just how nonlinear health outcomes are when trying to manage the negative impact that an environmental force like sprawl has on poor communities; $C6$ and $C8$ reminded scholars that power-relations, subjectivity, voice, and history are important dimensions of complex systems and need to be studied, whatever the topic; and $C9$ validated that studying agency adds an

important tool for studying the role agency play in the dynamics of a complex systems issue like sprawl and health; plus, it hints at the potential of agent-based modeling for conducting thought experiments and theoretical validation.

In terms of the negative test of the DTCS, the COP definition, overall, fit well with our case study, both empirically and theoretically. There is, however, a key limitation to the COP definition: it does not conceptualize health or health care in complex systems terms. Health is primarily treated as an outcome variable, rather than a system unto itself. Furthermore, health care is not addressed as a complex sub-system within the larger system of place.

Finally, in terms of our attempt to advance Keshavarz et al [38], our study tried to make two important points. First, if the application of complexity science to the study of place is to advance, scholars need to be more theoretically systematic in their usage of these ideas. The DTCS is an attempt to move scholars in this direction, offering a synthetically useful tool for testing the empirical validity and theoret-

ical value of defining places as complex systems. The second point is that community health scholars desperately need to advance the methods they use, catching up with those widely employed by complexity scientists in other disciplines. As an attempt to advance the field in this direction, we employed the SACS Toolkit, which we hope, demonstrated (a) how case-based modeling is an improvement over linear modeling and (b) how computational techniques, such as artificial intelligence (the SOM), cluster analysis, agent-based modeling, and network analysis can be used together to create new models of places as complex systems. However, as with all research, additional testing is necessary.

References

1. Aldenderfer, M.S., Blashfield, R.K.: Cluster Analysis. Sage Publications, Beverly Hills, CA (1984)
2. Auchincloss, A., Diez-Roux, A.V.: A new tool for epidemiology: Theusefulness of dynamic-agent models in understanding place effects on health. American Journal of Epidemiology **168**, 1–8 (2008)
3. Baao Fernando, V.L., Painho, M.: Self-organizing maps assubstitutes for k-means clustering. ICCS 2005,LNCS **3516**, 476–483 (2005)
4. Bar-Yam, Y.: Dynamics of Complex Systems. Westview Press, New York, NY (2003)
5. Bernard, P., Charafeddine, R., Frohlich, K.L., Daniel, M., Kestens, Y., Potvin, L.: Health inequalities and place: A theoretical conception of neighborhood. Social Science and Medicine **65(9)**, 1839–1852 (2007)
6. Blackman, T.: Placing health: Neighbourhood renewal, health improvementand complexity. The Policy Press **UK** (2006)
7. Bowles, S., Durlaf, S., Hoff, K.: Poverty Traps. Princeton University Press, Princeton, New Jersey (2006)
8. Browning, C., Bjornstrom, E., Cagney, K.: Health and mortality consequences of the physicalenvironment. Springer (2011)
9. Bruch, E., Mare, R.: Neighborhood choice and neighborhood change. American Journal ofSociology **112(3)**, 667–709 (2006)
10. Byrne, D.: Complexity Theory and the Social Sciences. Routledge, London, UK (1998)
11. Byrne, D.: Complexity, configurations and cases. Theory, Culture and Society. **22(5)**, 95–111 (2005)
12. Byrne, D.: Complex realist and configurational approaches to cases: A radical synthesis. In Byrne and Ragins The sage handbook of case-based methods. Sage Press, CA (2009)
13. Byrne, D., Ragin, C.: The sage handbook of case-based methods. Sage Press, CA (2009)
14. Capra, F.: The Web of Life. Anchor Books Doubleday, New York, NY (1996)
15. Castellani, B., F., H.: Sociology and complexity science: A new field of inquiry. Springer, Berlin, Germany (2009)
16. Castellani, B., Rajaram, R.: Case-based modeling and the sacs toolkit: A mathematical outline. Computational and Organizational Theory **to appear** (2012)
17. Cilliers, P.: Complexity and Postmodernism: Understanding Complex Systems. Routledge, New York, NY (1998)
18. Cox, M., Boyle, P.J., Davey, P.G., Feng, Z., Morris, A.D.: Locality deprivation and type 2 diabetes incidence: A local test of relative inequalities. Social Science and Medicine **65**, 1953–1964 (2007)
19. Cummins, S., Curtis, S., Diez-Roux, A.V., Macintyre, S.: Understanding and representing place in health research: A relational approach. Social Science and Medicine **65**, 1825–1838 (2007)
20. Curtis, S., M., S., Quesnel-Vallee, A.: Socio-geographic mobility and health status: A longitudinal analysis using the national population health survey of canada. Social Science and Medicine **69**, 1845–1853 (2009)
21. Curtis, S., Rees-Jones, I.: Is there a place for geography in the analysis ofhealth inequality? Sociology of Health and Illness **20**, 645–672 (1998)
22. Curtis, S., Riva, M.: Health geographies i: Complexity theory andhuman health. Progress in Human Geography **a**, 1–9 (2009)
23. Curtis, S., Riva, M.: Health geographies ii: Complexity and healthcare systems and policy. Progress in Human Geography **b**, 1–8 (2009)
24. Du, K.L.: Clustering: A neural network approach. Neural Networks **22**, 89–107 (2010)
25. Dunn, J., Cummins, S.: Placing health in context [editorial]. Social Science and Medicine **65**, 1821–1824 (2007)
26. Durie, R., Wyatt, K.: New communities, new relations: The impactof community organization on health outcomes. Social Science and Medicine **9**, 1928–1941 (2007)
27. Epstein, J.: Generative Social Science: Studies in Agent-Based ComputationalModeling. Princeton University Press, Princeton, NJ (2007)
28. Gatrell, A.C.: Complexity theory and geographies of health: A criticalassessment. Social Science and Medicine **60**, 2661–2671 (2005)
29. Gilbert, N., Troitzsch, K.: Simulation for Social Scientists, 2nd Edition. Open University Press., Buckingham, PA (2005)
30. Gladwell, M.: The tipping point : how little things can make a big difference. Little, Brown, Boston, MA (2000)
31. Gleick, J.: Chaos: Making a New Science. Penguin Books, New York, NY (1987)
32. Haggis, T.: Approaching complexity: A commentary on keshavarz, nutbeam, rowling andkhavarpour. Social Science and Medicine **70**, 1475–1477 (2010)
33. Hammond, D.: The Science of Synthesis: Exploring the Social Implications of General Systems Theory. University Press of Colorado, Colorado (2003)
34. Holland, J.: Emergence: From Chaos to Order. Perseus Books, Cambridge, MA (1998)
35. Jain, A.: Data clustering: 50 years beyond k-means. Pattern Recognition Letters **31(8)**, 651–666 (2010)
36. Kamada, T., Kawai, S.: An algorithm for drawing general undirected graphs. Information Processing Letters **31**, 7–15 (1989)
37. Kauff man, S.A.: Investigations. Oxford University Press (2000)
38. Keshavarz, N., Nutbeam, D., Rowling, L., Khavarpour, F.: Schools as social complex adaptive systems: A new way to understand the challenges of introducing the health promoting schools concept. Social Science and Medicine **70**, 1467–1474 (2010)

39. Klir, G.J.: Facets of Systems Science, 2nd edition. KluwerAcademic/Plenum Publishers, New York, NY (2001)

40. Kluver, J., C, K.: Social Understanding: On Hermeneutics, Geographical Models and Artificial Intelligence. Springer, Germany (2011)

41. Kohonen, T.: Self-Organizing Maps, 3rd Edition. Springer, New York, NY (2001)

42. Kuo, R.J., Ho, L.M., , Hu, C.M.: Integration of self-organizing feature map andk-means algorithm for market segmentation. Computers and OperationsResearch **29**, 1475–1493 (2002)

43. Lewin, R.: Complexity: Life at the edge of chaos. University of Chicago Press, Chicago, IL (1992)

44. Lin-Fisher, B.: Forum fighting hospital plan: Akron general hosts even thatargues against doctor-owned facility in northern summit. Akron Beacon Journal **10-22-2008** (2008)

45. Macintyre, S., Ellaway, A., Cummins, S.: Place effects on health:how can we conceptualise, operationalise and measure them? Social Science and Medicine **55**, 125–139 (2002)

46. Maturana, H., Varela, F.: The Tree of Knowledge: The Biological Roots ofHuman Understanding. Shambala, Boston, MA (1998)

47. Mitchell, M.: Complexity: A Guided Tour. Oxford University Press (2009)

48. Morrison, K.: Complexity theory, school leadership and management:questions for theory and practice. Educational Management Administration and Leadership **38(3)**, 374–393 (2010)

49. Peterson, D., Brody, T., Davis, W., Goranson, S., Lafaire, M., Lee, J., Lehrman, L., Lutner, L., Zeller, A.: Urban sprawl modeling, air quality monitoring, and risk communication: The northeast ohio project. Environmental Protection Agency **EPA/625/R-02/016** (2002)

50. Pickett, K.E., Pearl, M.: Multilevel analysis of neighbourhood socioeconomiccontext and health outcomes: A critical review. Journal of Epidemiology and Community Health **55**, 111–122 (2001)

51. Rihoux, B., Ragin, C.: Configurational ComparativeMethods: Qualitative Comparative ANalsysi (QCA) and Related Techniques. Sage (2009)

52. Riva, M., Curtis, S., Gauvin, L., Fagg, J.: Unravelling the extent of inequalities in health acrossurban and rural areas: evidence from a national sample in england. Social Science and Medicine **68**, 654–663 (2009)

53. Robert, S.: Socioeconomic position and health: 'the independent contribution ofcommunity socioeconomic context'. Annual Review of Sociology **25**, 489–516 (1999)

54. Sampson, R., Morenoff, J., Gannon-Rowley, T.: Assessing 'neigborhoodeffects': social processes and new directions in research. Annual Review of Sociology **28**, 443–478 (2002)

55. Smith, J., Jenks, C.: Qualitative Complexity. Routledge, UK (2006)

56. Smith, K.P., A., C.N.: Social networks and health. Annual Review of Sociology **34**, 405–429 (2008)

57. Sridharan, S., Tunstall, H., Lawder, R., Mitchell, R.: An exploratory spatial data analysis approach to understanding the relationship between deprivation and mortality in scotland. Social Science and Medicine **65(9)**, 1942–1952 (2007)

58. Summit, H.: Healthy summit 2000 health indicators summary report. URL http://www.healthysummit.org/HS2010/ HSprjreports.cfm

59. Urry, J.: Global Complexity. Blackwell Publishing, Oxford, UK (2003)

60. Waldrop, M.M.: Complexity: The emerging science at the edge of order and chaos. Simon and Schuster paperbacks, New York, NY (1992)