

Cleveland State University
EngagedScholarship@CSU



Physics Faculty Publications

Physics Department

6-1-2007

Urban Spatial Pattern as Self-Organizing System: An Empirical Evaluation of Firm Location Decisions in Cleveland–Akron PMSA, Ohio

Mukesh Kumar

Jackson State University, mukesh.kumar@jsums.edu

William M. Bowen

Cleveland State University, bowen.urban@csuohio.edu

Miron Kaufman

Cleveland State University, m.kaufman@csuohio.edu

Follow this and additional works at: https://engagedscholarship.csuohio.edu/sciphysics_facpub

 Part of the [Physics Commons](#)

How does access to this work benefit you? Let us know!

Publisher's Statement

The final publication is available at Springer via <http://dx.doi.org/10.1007/S00168-006-0097-Z>

Repository Citation

Kumar, Mukesh; Bowen, William M.; and Kaufman, Miron, "Urban Spatial Pattern as Self-Organizing System: An Empirical Evaluation of Firm Location Decisions in Cleveland–Akron PMSA, Ohio" (2007). *Physics Faculty Publications*. 240.
https://engagedscholarship.csuohio.edu/sciphysics_facpub/240

This Article is brought to you for free and open access by the Physics Department at EngagedScholarship@CSU. It has been accepted for inclusion in Physics Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.

Urban spatial pattern as self-organizing system: An empirical evaluation of firm location decisions in Cleveland–Akron PMSA, Ohio

**Mukesh Kumar William M. Bowen
Miron Kaufman**

Abstract Economic models of urban spatial patterns have largely ignored complexity as an attribute of urban systems. Complexity theorists on the other hand have not made sufficiently serious and sustained attempts to verify empirically the relevance of complex systems models for urban spatial patterns. This research bridges this gap by simulating the evolution of an urban employment pattern as a self-organizing complex system and seeking its empirical validation. It estimates the model's parameters by using firm data aggregated to the level of municipalities in Cleveland-Akron Consolidated Metropolitan Statistical Area in Ohio. The interaction among four parameters, forces of attraction and dispersion and their respective rates of dissipation with distance, are modeled as a two-dimensional complex system. The research compares the states of the modeled system with empirical data to present viable methods for verification, calibration and validation of such models.

Introduction

Some of the most distinctive aspects of the modern human settlement pattern are seen in the birth and development of urban areas, most of which contain one or more cities. In the past, it was the norm to find a central location with concentrated employment at the core of an urban area surrounded by mostly residential developments. Most residents lived in either the central city or the surrounding residential neighborhoods and found employment in the core. Thus while the employment pattern was relatively concentrated, the residential pattern was more widely distributed.

The image of this “mono-centric” city has influenced a substantial amount of research in several disciplines including regional science, economics, geography, sociology, history and political science. The picture is one of skyscrapers in the center of the urban area surrounded by concentric rings of commerce, industry and low-lying residential developments. The core is seen as being well connected to the periphery by transportation and communication networks. The cost of interaction between core and periphery is typically viewed as an increasing function of distance from the core. This image is even today dominant in most of the research on urban spatial patterns, irrespective of its disciplinary origins.

While the commonly shared abstraction of a mono-centric city in many respects transcends disciplines, the disciplinary flavors in most of the research are impossible to miss. For instance, while economists tend to focus on theories of agglomeration, geographers emphasize the spatial aspects of settlements, sociologists study neighborhood dynamics, and historians argue about definite events that shaped the current patterns. Ironically, although all of this research has over the years led to great strides in understanding urban spatial patterns, meanwhile ever-changing urban areas have metamorphosed into polycentric patterns. This, in turn, has rendered the mono-centric city image largely and increasingly implausible and uninformative. Some of the earlier insights and images have become practically obsolete and have brought about fresh sets of challenges.

One of these challenges is understanding polycentric urban spatial patterns. Accordingly, a relatively new approach—the “complex systems approach”—has gained some credence. This is in many respects a new paradigm that has emerged during the last quarter of the twentieth century. A broad outline can be garnered from several works by Allen (1997, 2001), Weidlich and Haag (1987), and White and Engelen (2000). Perhaps the most important feature is the treatment of equilibrium. The focus is on a multiplicity of attractors in a system, together with their related states and the stochastic considerations that would shift a system from one state to another. This focus leads to conceptual formulations such as self-organization, spontaneous order, and multiplicity and indeterminacy of equilibria. These formulations, however, have proven to be difficult to calibrate and therefore to test.

In spite of increasing interest in complexity, there is no single, generally accepted and testable theory in urban economics and elsewhere. In spite of some efforts to formulate such a theory (Kauffman 1993), currently there are

only fields and techniques loosely grouped around the banner of “Sciences of Complexity”. Their application to economic issues is characterized most importantly by two features: the standard neo-classical economic assumption of diminishing returns to scale is abandoned, and the possibility of increasing returns to scale is considered. Accordingly, an economic system is likely to acquire self-reinforcing tendencies.¹ Arthur (1988) notes that if an economic system exhibits such self-reinforcing tendencies, it will be characterized by one or more of the following four properties: multiple equilibria, possible inefficiency, path-dependence, and lock-in.² These properties are mostly used as classification schemes for complex systems.

This research develops an empirical test for a complex systems model proposed by Krugman (1996) that specifically focuses on the location of firms. Theoretical extensions of this approach to other aspects of urban spatial structure such as economic development, neighborhood dynamics, and housing patterns are feasible, though beyond the scope of the current research.

Background

In this section, a short literature review is followed by a brief introduction to Krugman’s model of firm location. It begins with a brief outline of the most significant and relevant aspects of the evolution of location theory from three distinct perspectives: the spatial perspective of the German tradition; the market perspective of neo-classical economics; and the more recent complex systems perspective.

Literature review

The research on the location of economic activity can be traced largely to three traditions — the land use and land rent analysis of Von Thünen (1826), followed by optimal plant location through Weberian (Weber 1909) triangles, and central place theory of Christaller (1933) and Lösch (1940). Alonso (1964) reinterpreted the von Thünen model by substituting commuters for farmers and a central business district for the isolated town. The Weberian tradition focused primarily on the optimization of location choices and has influenced most current transportation cost minimization models. Possibly the most important work in that tradition was Moses’ Location Invariance Theorem (MLIT)

(Moses 1958). Central place theory has inspired another set of literature that has dealt with urban hierarchies (Fujita et al. 1999).

Attempts to model the allocation of scarce resources among alternative and sometimes competing geographic locations form the core of location theory. The overarching concern in most of this literature remains focused upon how equilibrium comes about under different forms of competition. There are several excellent overviews available on the topic such as Anas et al. (1998), Kilkenney and Thisse (1999), Huriot and Thisse (2000), and Fujita and Thisse (2002).

If a much longer time-horizon is considered, the interdependency of choices made by firms and households becomes a dominant consideration. Thus the constraint of land availability becomes binding. Land values become a variable, which is usually endogenized. Transportation cost is no longer the dominant variable in determining land value because in the long run firms as well as households choose locations. Accordingly, Thünen-Alonso-type models of firm and household consumption of urban land were reformulated (Krugman, 1981; Helpman and Krugman, 1985; Fujita and Thisse, 1986). Firms were assumed to face a spatial distribution of demand established by the household consumption of urban land on one hand and to compete for locations on lines similar to as described by Hotelling (1929) on the other. One such formulation presents the interplay of centrifugal and centripetal forces, and is explicit about the interdependency of location choices made by firms/households. For households, these forces take the forms of attraction toward more differentiated locations and repulsion from higher rents in such locations. For retail firms, these forces of attraction are specifically attributable to higher concentration of consumers and the forces of repulsion from competition with other firms (Fujita and Thisse, 2002).

Currently, the dominant explanation for the emergence of employment centers tends to refer to the existence of external economies. More specifically, the explanation refers to three classic Marshallian sources of external economies: effects of market-size, thickness of labor markets, and pure external economies (Krugman 1998). Krugman (1998) calls them “centripetal forces.” At the same time, the forces of dispersion counter agglomeration. Krugman calls them “centrifugal forces.” Such forces have three components. First, dispersed immobile factors can lead to dispersed production (supply side) and dispersed production can lead to dispersed markets (demand side). Accordingly, some production will take place either close to its factors or close to its consumers. Second, increase in the geographical concentration of economic activity will cause land-rents to increase, thereby providing a disincentive to further agglomerate. Finally, there are pure external diseconomies such as higher commuting costs due to congestion. On balance, the tension between the forces of agglomeration and the forces of dispersion determines the level of employment concentration.

But the big question remains whether the interaction between the forces of agglomeration and dispersion leads to a stable state that could approximate a sense of equilibrium. At this stage, a researcher could go in one of the two ways. One could either remain within the conventional equilibrium framework

or choose the approach of complex systems modeling.³ While Fujita and Thisse (2002) have remained within the equilibrium framework, Krugman (1996) in selecting the latter way made an attempt to model the emergence of employment centers as self-organizing complex systems.

Krugman's model

Krugman (1996) presents the relationship between forces of attraction and forces of dispersion for a one-dimensional metropolitan area. He begins by specifying the interdependence between firm locations, where x is some location on the line, and $\lambda(x)$ is the density of firms at that location. It is assumed that the desirability of any given location depends both positively and negatively on the density of firms at other locations. Both forces decline with distance. He presents the relationship in the form of market potential as

$$P(x) = \int_z \left[A \exp(-r_1 D_{xz}) - B \exp(-r_2 D_{xz}) \right] \lambda(z) dz \quad (2.2.1)$$

where D_{xz} is the distance between locations x and z , and where $r_1 > r_2$; “ A ” represents the strength of the forces of agglomeration (centripetal force) and “ B ” represents the strength of the forces of dispersion (centrifugal force); and r_1 and r_2 represent the rates at which these forces dissipate with distance. Accordingly, assuming that businesses will gradually move toward locations with above average market potential P_{ave} and away from those with below average market potential, the average market potential is defined as

$$P_{ave} = \int_x P(x) \lambda(x) dx \quad (2.2.2)$$

The decision rule by which firms move is given as

$$\frac{d\lambda(x)}{dt} = \gamma [P(x) - P_{ave}] \lambda(x) \quad (2.2.3)$$

$\lambda(x)$ in Eq. (2.2.3) serves two purposes. It ensures that sum of changes add to zero and it also ensures that density at any given location does not fall below

³ The simplifying assumptions made to induce tractability in equilibrium-based modeling are often—but not always—left unquestioned. The possibility that increasing returns can cause multiple equilibria and possible inefficiency was noted by Marshall, albeit in an appendix (1891, Appendix H). In the 1970s, Arrow and Hahn (1971) and Brown and Heal (1979) explained the possibility of increasing returns to scale and consequent multiple equilibria and inefficiency. Another analytical problem results from circular causation derived from the market-size effects. As Krugman (1991) explains it, *ceteris Paribus* a manufacturing firm will locate close to where the demand for its product is the highest. But the demand will be higher at locations with other manufacturing firms because some of the demand will come from within the manufacturing sector itself. Myrdal (1957) called this demand-reinforcing aspect of economic size “circular causation,” in which, the “backward linkages” of Hirshman (1958) are reinforced by the “forward linkages.”

zero. In his simple demonstration, Krugman begins with the initial distribution of uniformly dispersed firms on a circle and demonstrates regularities in resulting distribution for varying values of the parameters.

Empirical calibration of the model

The basic task in this research is to empirically calibrate Krugman's (1996) formulation in a real world setting by estimating four parameters in his model. The parameters are concepts of agglomeration and dispersion, and their respective decay with distance. To conduct an empirical investigation of Krugman's claims of having specified a process with which to model a self-organizing economy through the complex systems approach, this research operationalizes self-organization in terms of entropy. As a proxy for self-organization, entropy estimates the spatial disorder in the urban area. Using entropy to measure for self-organization it becomes possible to test for the model's ability to predict the future distribution of employment. The level of correct prediction will determine the strength of the model.

Research design

The design for the proposed research closely follows standard simulation or scenario methods. There are three main components of this design: extensive use of simulation modeling, with a well-defined observation protocol; use of both inductive and deductive reasoning; and, comparison of simulation experiments with actual data. Simulation modeling will involve running discrete event simulation models, which uses entropy measures to describe and distinguish urban configurations. The spatial organization of employment centers is mathematically represented within a defined region by a complex systems model in which polycentric employment nodes naturally emerge from a self-organizing system. States of the system are compared with empirical data to verify, calibrate, and test the model. Success achieved at model calibration and at explaining the spatial configurations produced by the model for relatively high predictive ability will determine its empirical relevance.

The theory of cellular automata guides the division of geographical space into cells. Each cell mimics a city within Cleveland-Akron CMSA. A dynamic variable is associated with each cell (number of firms). Since a cellular model in which changes in states are triggered by an agent's action and the interaction among agents, the object based approach to cell definition poses less difficulty in defining protocols for "state transitions." This will suggest that when meaningful cells are of paramount importance, as they are in this case, an irregular lattice of cellular divisions mimicking cities (administrative boundaries) is of greater value. Cities differ from one another in terms of the bundle of goods and services that they provide to firms and relative tax burden. In general, municipalities can also be differentiated in terms of intangible characteristics such as business environment. Such differentiability does not exist for other boundaries such as zip codes or census tracts. Therefore, cells are more meaningful when they

mimic cities as compared to other types of divisions. For the present research, it can be argued that using municipalities is useful for capturing “Tiebout effects” (Tiebout 1956). However, when meaning attached to cells is of less importance, the space could also be represented at higher resolution. A successfully implemented example of a model of free agents in cellular space is given in Portugali (2000). The use of non-uniform cells mimicking municipalities would constitute a major advance in the application of cellular models to urban spatial pattern.

The sketch of the model is given in Chart 1. The upper half of the model uses the empirical data and the lower half of the model uses simulation. This research uses the data for the year of 1989 as the initial year and follows the changes observed for the subsequent years of 1991, 1993, 1995, and 1997. In the initial year of 1989, both the actual distribution and the simulated distribution are the same. While the empirical distribution remains mere observations for subsequent years, the distribution for simulation model is determined by the four parameters that are allowed to take values between 0 and 1. The best set of parameter values are estimated for the period of 1989 to 1997. Using this set, the simulation model predicts the distribution for the year 2001. The predicted distribution is compared to actual distribution to judge the goodness of the model.

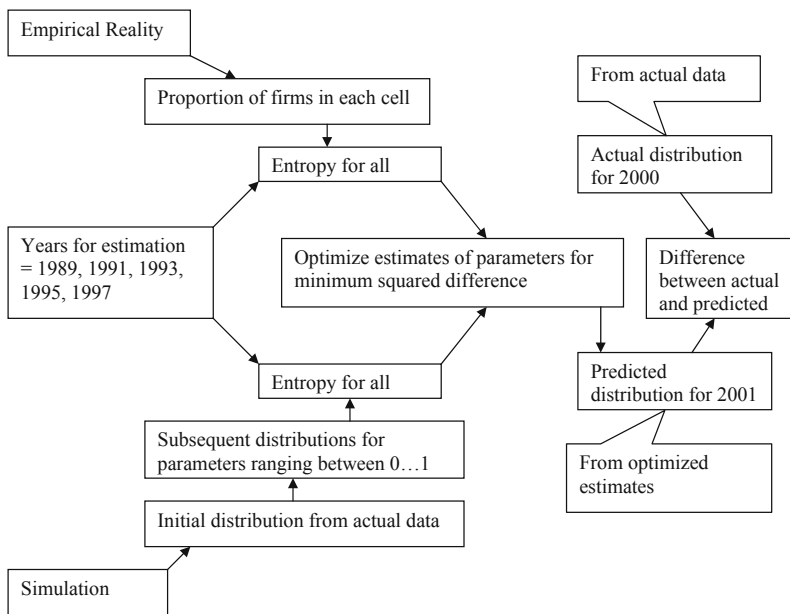


Chart 1 The model

A brief sketch of the process of empirical estimation of parameters, which, for a given initial distribution, determines the distribution of employment within a given spatial unit for a future time, is given in Chart 1.

We begin with the formulation for the distribution of employment within a spatial unit as given by Krugman. We discretize and simulate his formulation. The simulation model gives us predictions for subsequent time periods.

Let us suppose that we use an entropy number to represent the organization of firms as predicted by simulation. Based on Krugman's model, for a given initial distribution in time $t = 0$, we estimate the distribution of firms within the space by calculating the entropy for all subsequent times. If we represent entropy by " H " we call this $H(K_0)$ for the first period to $H(K_n)$ for the n th period. Now we calculate the actual entropy based on data for all the time periods. Let us call this $H(R_0) \dots H(R_n)$. We estimate the parameter values by using an exhaustive search method. We search for the set of estimates that yields the smallest sum of squared differences between simulated and actual entropies. We can numerically compute the four parameters for the closest possible equality between $H(K_1)$ and $H(R_1)$. Using this set of parameters, the research predicts the distribution for a given year in future. The accuracy of retroactive prediction of past distributions will determine the reliability of the estimates.⁴

Measuring entropy

The idea of entropy as a measure of geographic concentration of economic activity is not new. It has been widely used as a robust and rigorous measure of disorder or system homogeneity. Berry and Schwind (1969), Medvedkov (1967), Semple and Golledge (1970), Georgescu-Roegen (1971) and Garrison and Paulson (1973) were some of the early pioneers to use entropy as a measurement of the spatial concentration of economic activity. In most cases, it is used to describe the probability associated with locations with economic activity. The most popular derivation of entropy used in describing economic phenomenon is what is known as Shannon entropy, which was given by Shannon and Weaver (1949) in the context of communications theory. In its simplest form when applied to economic activity, it can be given by

$$H = - \sum_{i=1}^k P_i \ln(P_i) \quad (3.2.1)$$

where \ln denotes natural logarithms, and P_i denotes the relative share of economic activity in location i . If all the activity were located in only the j th location, since $\ln(P_j) = 0$ and all other $P_i = 0$, H will be equal to zero. At the other extreme, if economic activity were equally distributed among all k locations, H will be equal to $\ln(k)$. In order to make the entropy estimates vary between 0 and 1, we can use relative entropy G that can be given by

$$G = \frac{H}{\ln(k)} \quad (3.2.2)$$

In this research, since we use a non-uniform cellular structure, the calculation of relative entropy must be modified to account for the area of each cell (i.e. municipality). Estimating maximum entropy as following does this:

$$S_{\max} = \ln \left(\frac{\sum A}{A_{\min}} \right) \quad (3.2.3)$$

where A is the area. To account for the area in the entropy estimate we modify the formula as

$$S = - \sum_i P_i \ln \left(\frac{P_i}{A_i} \right) - \ln(A_{\min}) \quad (3.2.4)$$

The relative entropy now is given by

$$G = \frac{S}{S_{\max}} \quad (3.2.5)$$

Intuitively, in order to self-organize, an economy must observe a decrease in entropy. That is to say that entropy for any given initial distribution should decrease in subsequent years. However, we must use caution. In the simulations presented by Krugman in his book, the initial distribution of firms is uniform, which corresponds to a value of “1” for relative entropy. Therefore, unless the system remains frozen or cells gain or lose equal numbers of firms, any change will result in loss of entropy. But when we use actual data, the initial distribution at the time of the formation of the cities is often indefinite. If we use the distribution of a more recent year as the initial distribution, we do not have the option of beginning with an entropy value of “1.” Since the system is already in motion and the internal dynamics do not necessarily have the same beginning point as the available data, we cannot have the same expectation of decrease in entropy. The expectation of increase or a decrease in entropy should be based on the initial distribution used for the empirical investigation. This is especially important for detecting and observing the formation of edge city dynamics. If the initial distribution is dominated by one or two centers that gradually lose their dominance and yield some of the share of the economic activity to upcoming locations, the expectation will be reversed. An empirical test for self-organization is beyond the scope of this research and is potentially an important issue to explore in future research.

Study population and variables

The data used to calibrate the model are from the Cleveland-Akron Consolidated Metropolitan Statistical Area (CMSA), which is home to the city of Cleveland, OH. A map of the study area is included in Appendix 1. The selection of Cleveland-Akron CMSA is arbitrarily based most importantly on the reason that familiarity with the area helps in ensuring the quality of the

available data. More specifically, a much higher percentage of firms can be geocoded and cleaned to be included in the study. The CMSA has eight counties that have a total of 262 municipalities, of which Cleveland is the largest city followed by the city of Akron. Some of the municipalities did not have any firms in them. Data for some of the municipalities were unreliable. For example, data for municipalities with a single firm and no employment were considered unreliable. Such municipalities were dropped from the dataset used in the research. Finally, 192 municipalities were included for the models of all firms. For the model of non-population serving firms, 174 municipalities were included. For the model of population-serving firms, 155 municipalities were included.⁵ The distinction between population serving and non-population serving firms is based on the SIC codes. The codes of Division “G” (5211–5999), and Division “I” (7011–8999) were considered population serving. Division “G” refers to industries classified under “retail trade.” Division “I” refers to industries classified under “services.” This research assumes that the industries listed under retail and services sectors primarily serve the local population.⁶ All the others were considered non-population serving. There is only one variable in the model: density of firms in each municipality. The selection of firms⁷ instead of employment is based most importantly on three reasons. First, the operationalization of the decision-making process in terms of potential function for each municipality required that the decision-making unit could be located only at the firm level. Second, what makes a cell more or less attractive (centripetal and centrifugal forces) are realized at the firm level and not at the level of individual employment. In other words, agglomeration benefits and costs accrue to firms and not to individual employees. Third, while proposing the model being tested in this research, Krugman used firms as the decision-making unit. Since this research, more specifically, is an empirical investigation of Krugman’s model; it only makes sense to use firms in the model. For the purpose of this study cells have only two characteristics: their respective densities of firms and their relative location in terms of distance from other cells. Other characteristics such as zoning laws and transportation provision can be used for further refining the model in future research.

Data source

The data used in this research came from the ES 202 data. This dataset is based on employment, wage, and contributions data submitted to the Ohio

Table 1 Geocoding results

	1989	1991	1993	1995	1997	2001
Total records	63,922	64,479	71,689	72,179	75,825	77,682
Address listed in Ohio	58,966	59,381	67,742	68,829	72,451	74,321
Address in the CMSA	57,322	57,305	66,288	66,930	70,922	72,807
Geocoded	89.67%	88.87%	92.47%	92.73%	93.53%	93.72%

Department of Job and Family Services by employers subject to the Ohio Unemployment Compensation Law, as well as those covered under Unemployment Compensation for Federal Employees (UCFE). This dataset was available only for the period 1989–2001 at the time this research began. The spatial resolution for these years is quite high (point data at firm level). In an ideal situation, one would like to have a more exhaustive time period. However, to solve for the four unknown parameters, minimum required data is for five points in time. In this research, data for the years of 1989, 1991, 1993, 1995, and 1997 were used for estimating the parameters. With those parameters, the model predicted for year of 2001. The greatest difficulty with this data arises during the process of geocoding. Several of the records had their addresses listed outside the study area. Such addresses could not be included in the study. Table 1 lists the success rate after some painstaking effort.

The success rate for geocoding was worst for the year 1991 at 88.87% and the best for the year 2001 at 93.72%. Since the unmatched records cannot be tracked, one cannot make judgments about the biases or issues that could arise in any form. However, on a promising note, the collection of the ES 202 data has been getting better over the years. Future implementation with the same data will almost certainly have better geocoding rates.

Results

This section discusses how well a complex systems model of self-organizing economy does in describing the urban spatial pattern of the Cleveland-Akron CMSA. While the criterion for declaring the model “good” is subjective and simple, the process that it tries to describe is not. There are two criteria used for considering the model good or bad. First, how well does the model perform in predicting the overall spatial distribution of firms? Second, how well does the model predict number of firms in individual cells (municipalities in this research)?

Overall prediction

The Cleveland-Akron CMSA has a highly dispersed spatial distribution of firms. This observation is valid for all firms as well as population serving and

Table 2 Optimized parameter estimates

All industries						
Div(0-1)	Hours	A	B	Da	Db	Difference
8	43:06:47	1.000	0.750	1.000	0.250	0.000215
9	68:39:43	1.000	0.778	0.889	0.222	0.000220
10	83:56:00	1.000	0.800	1.000	0.300	0.000218
11	125:00:00	1.000	0.818	0.909	0.273	0.000212
12	217:00:00	1.000	0.750	1.000	0.250	0.000214
13	244:00:00	1.000	0.769	1.000	0.308	0.000216

non-population serving firms. The optimized values for all four parameters for different values of lambda are given in Table 2.

The actual entropy for the initial year of 1989 for all firms was 0.9020, which increased to 0.9136 for the prediction year of 2001. The expected relative entropy from the model was estimated at 0.9096 for a difference of 0.004. The actual entropy for the initial year of 1989 for population serving firms was 0.9102, which increased to 0.9239 for the prediction year of 2001. The expected entropy from the model was estimated at 0.9194 for a difference of 0.0045. The actual entropy for the initial year of 1989 for non-population serving firms was 0.9169, which increased to 0.9290 for the prediction year of 2001. The expected entropy from the model was estimated at 0.9261 for a difference of 0.0029. Table 3 lists the differences between actual and expected entropies.

Based on Table 3 one can conclude that the model predicts the overall spatial distribution of firms across the CMSA reasonably well. It works better for non-population serving firms than population-serving firms.

City-specific prediction

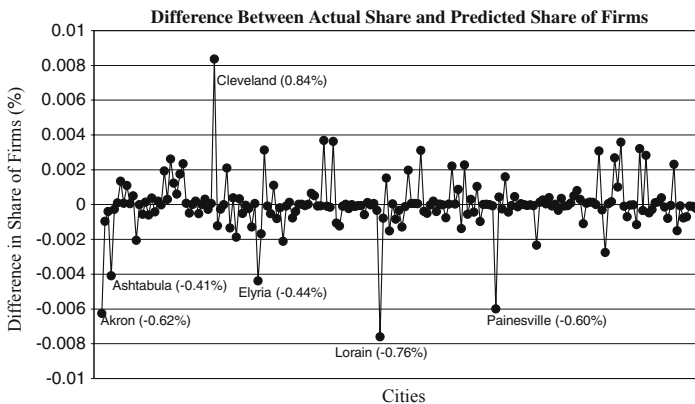
**Chart 2** City-specific prediction results for all firms

Table 3 Entropy results: actual and model

Type of industry	Actual	Model	Difference
All	0.9136	0.9096	0.0040
Population serving	0.9239	0.9194	0.0045
Non-population serving	0.9290	0.9261	0.0029

The predictive ability of this model to forecast the share of firms in each city seems promising. Chart 2 depicts the difference between actual share of firms across municipalities within the region for the year 2001 and the expected share of firms for all types. The predicted share of firms for all the cities is within the range of 1%. The worst predicted city is that of Cleveland, for which the model predicts to have 16.3 % of all firms to be located in. This is in comparison with 17.1% of the firms that were actually located in Cleveland for the year 2001. The next worst predicted city is that of Lorain. The model predicted it to have 2.3% of all firms. In reality, it had 1.6% of all firms. The third worst predicted city is that of Akron for which the model predicted 9.7%. This is in comparison with 9.1% of all firms that were located in Akron in 2001. There are only six cities for which the predicted share is off by greater than 0.4%. Aside from the three cities mentioned above, these include Painesville, Elyria, and Ashtabula. This is to suggest that of 192 cities under consideration, the model predicted the individual share of firms within the study area with the accuracy of $\pm 1\%$ for all the cities and with the accuracy of $\pm 0.4\%$ for 186 cities.

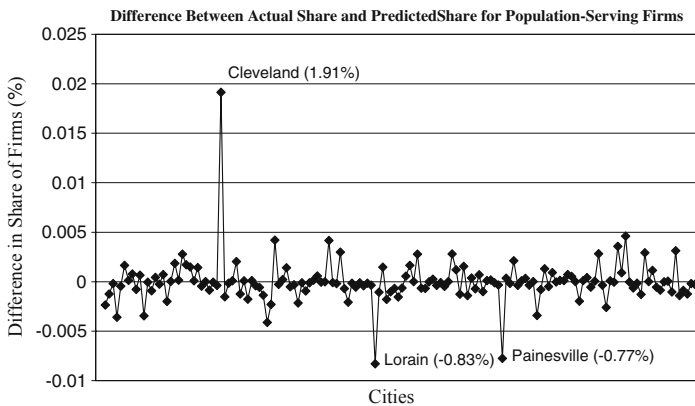


Chart 3 City-specific prediction results for population-serving firms

Chart 3 depicts the difference between actual share of firms within the region for the year 2001 and the expected share of firms for all population-serving firms. The predicted share of firms for all the cities is within the range of 2%. The worst predicted city is that of Cleveland, for which the model predicts to have 15.66 % of all firms to be located in. This is in comparison with 17.58% of the firms that

were located in Cleveland for the year 2001. The next worst predicted city is that of Lorain. The model predicted it to have 2.57% of all firms. In reality, it had 1.74% of all firms. The third worst predicted city is that of Painesville for which the model predicted 1.95%. This is in comparison with 1.18% of all firms that were located in Painesville in 2001. These are the only cities for which the predicted share is off by greater than 0.5%. This is to suggest that of 155 cities under consideration, the model predicted the individual share of population-serving firms within the study area with the accuracy of $\pm 2\%$ for all the cities, with the accuracy of $\pm 1\%$ for 154 cities and with the accuracy of $\pm 0.5\%$ for 152 cities.

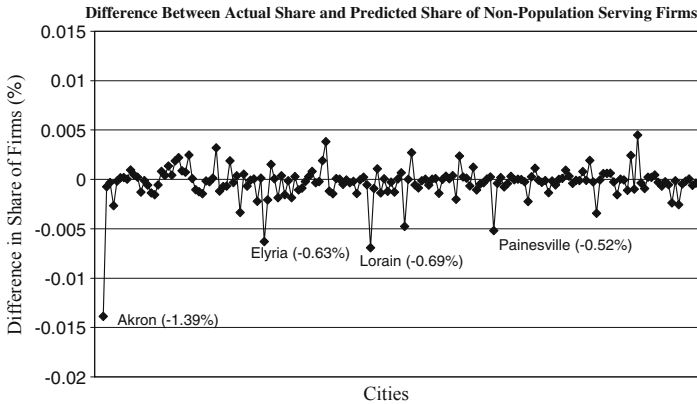


Chart 4 City-specific prediction results for non-population serving firms

Chart 4 depicts the difference between actual share of firms within the region for the year 2001 and the expected share of firms for all non-population serving firms. The predicted share of firms for all the cities is within the range of 1.5%. The worst predicted city is that of Akron, for which the model predicts to have 8.68% of all non-population serving firms to be located in. This is in comparison with 7.29% of the firms that were located in Akron for the year 2001. The next worst predicted city is that of Lorain. The model predicted it to have 1.93% of all firms. In reality, it had 1.24% of all firms. The third worst predicted city is that of Elyria for which the model predicted 2.62%. This is in comparison with 2% of all firms that were located in Elyria in 2001. The fourth worst predicted city is that of Painesville for which the model predicted 1.76%. This is in comparison with 1.24% of all firms that were located in Painesville in 2001. These are the only cities for which the predicted share is off by greater than 0.5%. This is to suggest that of 174 cities under consideration, the model predicted the individual share of population-serving firms within the study area with the accuracy of $\pm 1.5\%$ for all the cities, with the accuracy of $\pm 1\%$ for 170 cities and with the accuracy of $\pm 0.5\%$ for 169 cities.⁸

Discussion and conclusions

Undoubtedly the most theoretically well-developed approach to research on urban socio-spatial structure is based upon neo-classical urban economics. Such theory suggests that intra-urban spatial patterns of employment emerge through decisions made by firms largely in an effort to minimize the cost of inputs into their production processes in varying proportions at different locations. The relative prices of capital, land, and transportation, among other factors, thus play a crucial role in any such research into spatial patterns of firm location. However, while such theory is insightful and realistic in several important respects, in application it must either pre-specify the locations of employment centers without postulating the conditions under which they emerge, or else postulate the conditions under which employment centers emerge without explaining where those centers will be located (Cadwallader 1996, p. 66). Moreover, such theory assumes that all urban areas are monocentric.

Given that urban areas today tend more and more to be multiple-centered, the “multiple nuclei model” (Harris and Ullman 1945) may in this crucial respect come closer to qualifying as a realistic theory of urban spatial pattern than any of the mono-centric models. Its major feature is recognition that urban areas contain a number of centers or subcenters, rather than only one downtown, around which the patterns of employment and settlement are formed. However, that model is not specified with anywhere near the precision required to quantitatively predict employment centers or other features of urban spatial pattern. In fact, there exists broad agreement that the multiple nuclei model is far more of a descriptive rather than a predictive or explanatory device.

It seems that no perspective on urban spatial pattern will be fully coherent without a model that simultaneously specifies the parameters on polycentric employment centers and accounts for where such employment centers are located. Krugman’s model does both. Our research demonstrates that this model accounts for the relative spatial position and magnitude of employment centers within an urban economic system.

In light of the many polycentric urban areas observed on the landscape today, it furthermore seems that some level of re-description is required of the processes through which urban spatial patterns emerge. The emergence of polycentric urban spatial patterns is without a doubt in many respects dependent upon the same price factors as neoclassical urban economic theory stipulates— such as those related to capital, land, and transportation. But the interdependence between these factors seems to create complexity that forms and characterizes the larger context within which human purposes shape and

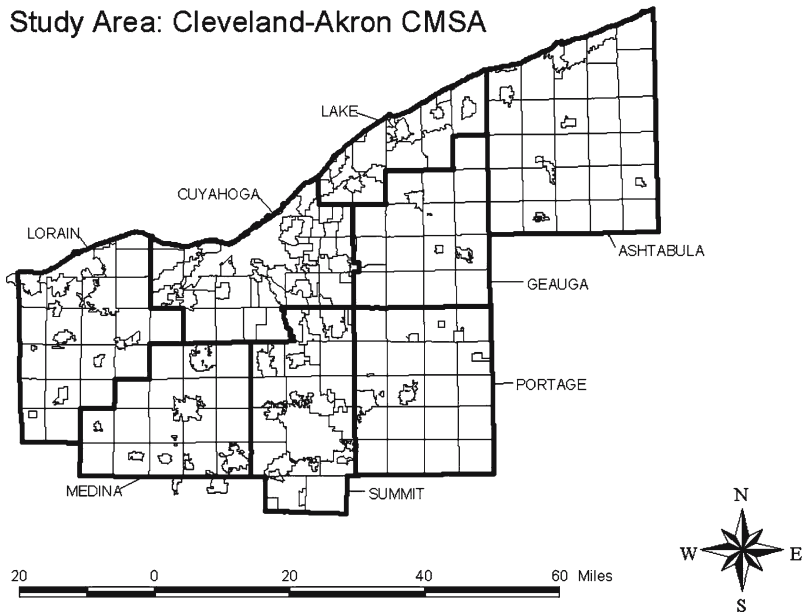
define boundary conditions on the pertinent social-spatial and economic systems. Thus it seems that a polycentric model with sufficient specificity to quantitatively predict outcomes is made coherent within the context of a complex systems approach.

In our view, the complexity created largely by interdependence between location factors leads to underconceptualizations that are partially responsible for firm location decisions. Moreover, a significant part of the challenge this presents is that the human mind does not have the power or means to measure with any precision whether or to what degree in any given decision situation it is engaged with conceptualizations involving complexity. From the point of view of the firm, this challenge is especially relevant in the process of describing the decision problem they face. That is, prior to selecting a location, the firm must at least specify the alternative possible locations, identify the cost factors associated with each of them, and evaluate how satisfactory or agreeable each location is in terms of these factors. In as much as models of firm location are based upon the assumption of rational processes of deductive reasoning, perfect information, a fixed and consistent preference ordering, and/or perfect cognitive processing power, they tend to ignore this challenge. Before a firm makes a choice of location, the relevant facts and circumstances must be depicted, characterized, and specified. This is the larger context in which complexity appertains. Only after the firm's location decisions are made does the discipline of the marketplace check bad decisions.

Models of any reasonably complex system must balance parsimony with the requisite variety of variables. Parsimony is necessary to facilitate understanding. The requisite variety of variables is required to facilitate accurate understanding. Numerous complex systems models have been developed in recent years that attempt to achieve parsimony by reference to simplifying assumptions about the character of agents making decisions. Complexity is found in the variables that reflect physical, economic, or natural systems, or else in the mathematical equations used to describe such systems. But in this paper we take an alternative view in which the model itself is parsimonious—it contains only four parameters—and the complexity is found within the incapacity of the human mind to retain and process the requisite variety of variables required to consistently select optimal locations. Moreover, this alternative view is the key to coherence in an otherwise powerful empirical model of polycentric firm location.

Appendix 1. Study area

Study Area: Cleveland-Akron CMSA



References

- Allen MP (1997) *Cities and regions as self-organizing systems: Models of complexity*. Gordon and Breach, London
- Allen MP (2001) What is the science of complexity? *Emergence* 3:24–42
- Alonso W (1964) *Location and land use*. Harvard University Press, Cambridge, MA
- Anas A, Arnott R, Small KA (1998) Urban spatial structure. *J Econ Literat* 36:1426–1464
- Arrow K, Hahn F (1971) *General competitive analysis*. Holden-Day, New York
- Arthur WB (1988) Self-reinforcing mechanisms in economics. In: Philip WA, Kenneth JA, David Pines (eds) *The economy as an evolving complex system*. Addison-Wesley, Reading, MA, pp 9–31
- Berry BJJ, Schwind PJ (1969) Information and entropy in migrant flows. *Geograph Anal* 1:5–14
- Brown D, Heal G (1979) Equity, efficiency and increasing returns. *Rev Econ Stud* 46:571–585
- Cadwallader M (1996) *Urban geography: An analytical approach*. Prentice-Hall, Upper Saddle River, NJ
- Christaller W (1933) *Die Zentralen Orte in Seddeutschland*. Translated by C. W. Baskin to *Central places in southern Germany*. Prentice-Hall, Englewood Cliffs, NJ, 1966
- Fujita M, Krugman P, Mori T (1999) On the evolution of hierarchical urban systems. *Eur Econ Rev* 43:209–251
- Fujita M, Thisse JF (1986) Spatial competition with a land market: hotelling and Von Thunen unified. *Rev Econ Stud* 53:819–841
- Fujita M, Thisse JF (2002) *Economics of agglomeration*. Cambridge University Press, Cambridge, UK
- Garrison CB, Paulson AS (1973) An entropy measure of the geographic concentration of economic activity. *Econ Geograph* 49:319–324
- Georgescu-Rogen N (1971) *The entropy law and economic process*. Harvard University Press, Cambridge, MA
- Harris C, Ullman EL (1945) The nature of cities. *Ann Ame Acad Polit Soc Sci* 242:7–17

- Helpman E, Krugman P (1985) *Market structure and foreign trade: Increasing returns, imperfect competition and the international economy*. MIT Press, London
- Hirshman OA (1958) *The strategy of economic development*. Yale University Press, New Haven, CT
- Hotelling H (1929) Stability in competition. *Econ J* 39:41–47
- Huriot JM, Thisse JF (ed) (2000) *Economics of cities*. Cambridge University Press, Cambridge, UK
- Kauffman AS (1993) *The origins of order*. Oxford University Press, New York
- Kilkenny M, Thisse JF (1999) Economics of location: a selective survey. *Comput Oper Res* 26:1369–1394
- Krugman P (1981) Trade, accumulation and uneven development. *J Develop Econ* 8:149–61
- Krugman P (1991) Increasing returns and economic geography. *J Polit Econ* 99:483–499
- Krugman P (1996) *The self-organizing economy*. Blackwell, Cambridge, MA
- Krugman P (1998) What's new about the new economic geography? *Oxford Rev Econ Policy* 14:7–17
- Lösch A (1940) *The economics of location* (Fischer Jena: English translation). Yale University Press, New Haven, CT (1954)
- Marshall A (1920) *Principles of economics*, 8th ed. McMillan, London (1st ed. [1891] McMillan, London)
- Medvedkov YV (1967) The concept of entropy in settlement pattern analysis. *Pap Reg Sci Assoc* 18:165–168
- Moses NL (1958) Location and the theory of production. *Quart J Econ* 72:259–272
- Myrdal G (1957) *Economic theory and under-developed regions*. Duckworth, London
- Portugali J (2000) *Self-organization and the city*, Springer, Berlin Heidelberg New York
- Simple RK, Golledge RG (1970) An analysis of entropy changes in a settlement pattern over time. *Econ Geograph* 46:157–160
- Shannon CE, Weaver W (1949) *The mathematical theory of communication*. The University of Illinois Press, Urbana, IL
- Tiebout C (1956) A pure theory of local expenditures. *J Polit Econ* 64:416–424
- Von Thünen JH (1826) *Der Isolierte Staat in Beziehung auf Landschaft und Nationalökonomie*, (translation by C. M. Wartenberg to von Thünen's *Isolated State*, 1966). Pergamon Press, Oxford
- Weber A (1909) *Über den Standort der Industrien* (translation: *The theory of the location of industries*, 1929). Chicago University Press, Chicago
- Weidlich W, Haag G (1987) A dynamic phase transition model for spatial agglomeration processes. *J Reg Sci* 27:529–569
- White R, Engelen G (2000) High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Comput Environ Urban Syst* 24:383–400