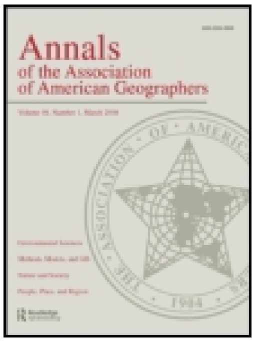

Annals of the Association of American Geographers

ISSN: 0004-5608 (Print) 1467-8306 (Online) Journal homepage: http://www.tandfonline.com/loi/raag20

A LOCATION THEORY FOR RURAL SETTLEMENT

JOHN C. HUDSON

To cite this article: JOHN C. HUDSON (1969) A LOCATION THEORY FOR RURAL SETTLEMENT, Annals of the Association of American Geographers, 59:2, 365-381, DOI: 10.1111/j.1467-8306.1969.tb00676.x

To link to this article: https://doi.org/10.1111/j.1467-8306.1969.tb00676.x

A LOCATION THEORY FOR RURAL SETTLEMENT[1]

JOHN C. HUDSON
University of North Dakota, Grand Forks

ABSTRACT. A theory of rural settlement location is proposed which will explain changes in settlement distribution over time. A series of spatial processes similar to those found in plant ecology studies are postulated for rural settlement. There are three phases: Colonization, by which the occupied territory of a population expands; spread, through which settlement density increases with a tendency to short distance dispersal; and competition, the process which produces a regularity in settlement pattern when rural dwellers are found in sufficient numbers to compete for space. Empirical investigations over a ninety-year period (1870–1960) in six Iowa counties reveals that the expected increase in regularity does occur. These effects are measured by fitting the Poisson, negative binomial, and regular Poisson distributions to quadrat censuses of the settlement maps. Variance-mean ratios declined over time with changes in the farm economy, requiring fewer, but larger farms. The negative binomial fit the early, more clustered distributions best, whereas the regular Poisson series fit the recent data best.

THE study of the distribution of urban and rural settlement has occupied an important position in the historical development of geography. Other subdivisions of the field such as urban geography, cultural geography, and agricultural geography, have recently taken over much of the subject matter once considered to be in the category of settlement. From reading Stone's 1965 review of definitive works on settlement, as well as other recent studies, it appears that rural settlement location, unlike urban settlement, is a subject on which little theoretical work has been done.[2] The purpose of this paper is to construct a theory of rural settlement location.

The theory will make use of several existing laws of settlement location, putting them together in such a way as to offer an explanation for the changes in settlement pattern that take place as a population invades, permanently occupies, and makes a living in an area. The focus is on agricultural settlement although the laws involved are sufficiently broad in scope that they may be applied to many geographical problems involving the arrangement of points with respect to each other, and with respect to the area in which they are located.

FORM AND PROCESS IN SETTLEMENT THEORY

At present, settlement geography is essentially concerned with the process of land settlement and the form which this settlement takes. The spatial distributions that are analyzed are usually those of dwelling places.[3] The terms form and process are apposite to geography's two primitives, respectively, geometry and movement.[4] Settlement form is the spatial realization of transformations on the population produced by settlement process, or settling.

Much attention has been paid to settlement form, especially by European geographers.[5] The easy conceptualization of a distribution of dwellings as a pattern of points on a map

Accepted for publication February 20, 1968.

[1] This article is a revised version of Chapter Two of the author's unpublished doctoral dissertation, *Theoretical Settlement Geography* (Iowa City: Department of Geography, University of Iowa, 1967).

[2] K. H. Stone, "The Development of a Focus for the Geography of Settlement," *Economic Geography*, Vol. 41 (1965), pp. 346–55.

[3] Stone, *op. cit.*, footnote 2, p. 347.

[4] W. W. Bunge, *Theoretical Geography* (Lund, Sweden: C. W. K. Gleerup, 1966), second edition, p. 210.

[5] A. Demangeon, "Une Carte de l'Habitat," *Annales de Geographie*, Vol. 42 (1963), pp. 225–32; M. A. Lefevre, *Principes et Problemes de Geographie Humaine* (Brussels: Editorial Office, 1945); J. M. Houston, *A Social Geography of Europe* (London: Duckworth, 1953), Ch. 5; J. A. Barnes and A. H. Robinson, "A New Method for the Representation of Dispersed Rural Population," *Geographical Review*, Vol. 30 (1940), pp. 134–37.

has no doubt facilitated this. Just as early geographic studies of shape were done by political geographers interested in boundaries, so were some of the first attempts to classify patterns done by settlement geographers. The terms linear, clustered, agglomerated, dispersed, and others which have re-entered geography via statistical methods, were used widely by settlement geographers in earlier studies. However, research has not progressed much beyond the stage of classification. Settlement process has recently been studied via simulation techniques.[6] The consideration of settlement spread as a diffusion process has promoted this approach to the construction of settlement patterns, using models which emulate the actual process of settling the land.

Nevertheless, statistical technique is no substitute for theory. The usefulness of theory and predictive models in geography is now a matter of record. Theories of location explain the laws of spatial distributions. Unless geographical explanations (or predictions) have a theoretical justification, the recognition of spatial regularities is of little value. There are four sources of location theory employed in this work. They are: central place theory, diffusion theory, ecological distribution theory, and a fourth category of generalizations encompassing aspects of all these three, morphological laws.

Probably the most developed theory of settlement in geography is central place theory. Although its urban aspects are most often studied, Lösch treated rural settlement location in his version of the theory.[7] As important as the theory of rural settlement in central place work, however, is the assumption that farms are uniformly distributed. One of the tasks of this paper is to examine the conditions on the validity of this assumption, both theoretically and empirically. Diffusion theory has been applied to settlement locations by Bylund.[8] His focus was on movement, or process, rather than form. A second task of this paper is to integrate diffusion theory with central place theory, as it applies to rural settlement. A guide as to how this may be done is offered by the plant and animal ecologists, whose location theories have considered form and process simultaneously. Although the nonspatial aspects of ecological theories are quite unlike human location theory, the spatial properties are similar. The justification for this borrowing is provided by the last category of location theories cited above, morphological laws. Such laws are concerned with forms in a generic sense and frequently involve extremum problems, asking to maximize or minimize some property of a mapped distribution.[9] Morphological laws make scant mention of the phenomena involved, and concentrate solely on spatial properties.

THE PROCESSES OF RURAL SETTLEMENT

In the spatial distribution theory of plant and animal ecology several spatial processes are apparent.[10] A series of similar processes is postulated for rural settlement. Although the processes are not always clearly identified it seems that it is characteristic of a distribution to pass through three phases:

1) A phase of colonization occurs; the species invades a new area, extending its habitat beyond the borders of its former environment;

2) biological renewal produces a regeneration of the species through an increase in numbers with a general tendency to short-distance dispersal, filling up the gaps in the distribution formed by the original colonizers, and as time passes the process is checked by a third set of forces;

---

[6] P. Haggett, *Locational Analysis in Human Geography* (New York: St. Martins Press, 1966), pp. 87–100.

[7] A. Lösch, *The Economics of Location* (New Haven: Yale University Press, 1954).

[8] E. Bylund, "Theoretical Considerations Regarding the Distribution of Settlement in Inner Northern Sweden," *Geografiska Annaler*, Vol. 62 (1960), pp. 225–31.

[9] Bunge, *op. cit.*, footnote 4, pp. 249–50.

[10] A few examples of articles discussing spatial processes in ecology are: A. R. Clapham, "Over-dispersion in Grassland Communities and the Use of Statistical Methods in Plant Ecology," *Journal of Ecology*, Vol. 24 (1936), pp. 232–51; N. A. Holme, "Population Dispersion in *Tellina tenuis* da Costa," *Journal of Marine Biology*, Series A, Vol. 29 (1959), pp. 267–80; M. E. Phillips, "Studies in the Quantitative Morphology and Ecology of *Eriophorum angustifolium*. II: Competition and Dispersion," *Journal of Ecology*, Vol. 42 (1954), pp. 187–220; K. A. Kershaw, "Pattern in Vegetation and its Causality," *Ecology*, Vol. 44 (1963), pp. 377–88; A. M. Laessle, "Spacing and Competition in Natural Stands of Sand Pine," *Ecology*, Vol. 46 (1965), pp. 65–72.

3) owing to limitations of the environment, weak individuals are forced out by their stronger neighbors, density tends to decrease, and pattern stabilizes.

The first process is here termed *colonization*, that phase associated with the dispersal of settlement into new territory, or a new environment, or into an unoccupied portion of the old environment. The second process is termed *spread*. Characteristic of spread is increasing population density, creation of settlement clusters and eventual pressure on the environment, both physical and social. The third process, *competition*, is best documented in geographical location theory. It is this process that tends to produce great regularity in the settlement pattern and in turn produces one condition for the regular network of central places.

The three processes are, in a sense, actually two. Colonization and spread are both aspects of diffusion, but models of the two processes are more simply studied in separate models to reduce the number of parameters. Furthermore, each process has its own spatial properties, making separate treatment more valuable.

Colonization

The ensuing arguments concerning colonization rely heavily on definitions, terminology, and a theoretical treatment of the subject of animal territories found in an article by Hutchinson.[11] In the present work, a gradient function governing settlement density is introduced, which was not present in Hutchinson's model. Although the notation and terminology may at first seem unnecessarily formal, the use of the ecological model will simplify the discussion later when it is desired to enforce or relax various uniformity assumptions concerning the environment.

The existence, and magnitude of human settlement in an area may be thought of as being contingent upon m environmental variables, from which there may be derived a set of n variables which are independent in the statistical sense $n < m$. Thus, the set of all possible values that the n linearly independent variables can take forms an n-dimensional vector space. This is called the *niche space*, (N). Each vector in this space has n components, defining a certain combination of values of the environmental variables. It is a familiar fact that there exist vectors $(x_{i1}, x_{i2}, x_{i3}, \ldots x_{in})$ that are too extreme to permit human settlement. There exist values that are too cold, too warm, too wet, too dry, too far from markets, for settlement to exist, for a given level of the arts.[12] The bounded subset of points in this n-space that includes the permissible values is the innermost intersection of the subsets of permissible valuables on the n variables. This subset is an n-dimensional hypervolume, and is called the fundamental niche of the population (Fig. 1).

A change in technology may bring about a change in the permissible conditions under which settlement may take place, thus the size of the hypervolume is flexible. For example, commercial grain farming became feasible in the Great Plains with the coming of railroads, thus the transportation limit was eased permitting settlement in the area.

It is important to note that this N-space is not physical space, or a mapped abstraction of it. Ecologists use the term *biotope space*, (B), for physical space wherein the variables of the niche space are realized.[13] It is this space that geographers map. There exists a not necessarily single-valued function $\Phi(x_1, x_2, x_3, \ldots x_n) = B$ from the niche space into the biotope space, assigning characteristics to areas. Because settlement is feasible under a certain set of conditions is no reason to expect these conditions actually have physical expression in a given space, B. If no area has the characteristic combination then the image of the mapping is the null set.

Limits in the niche are counterparts of distributional boundaries (regions) in the biotope. An example of a limiting value on a variable in the niche space is minimum farm size—the size beneath which agricultural operations are economically unfeasible. Tree line, on the other hand, is a boundary

[11] G. E. Hutchinson, "Concluding Remarks," *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. 22 (1957), pp. 415–27.

[12] This is directly analogous to the concept of "Optima and Limits," in H. H. McCarty and J. B. Lindberg, *A Preface to Economic Geography* (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1966), p. 220.

[13] Hutchinson, *op. cit.*, footnote 11, p. 418.

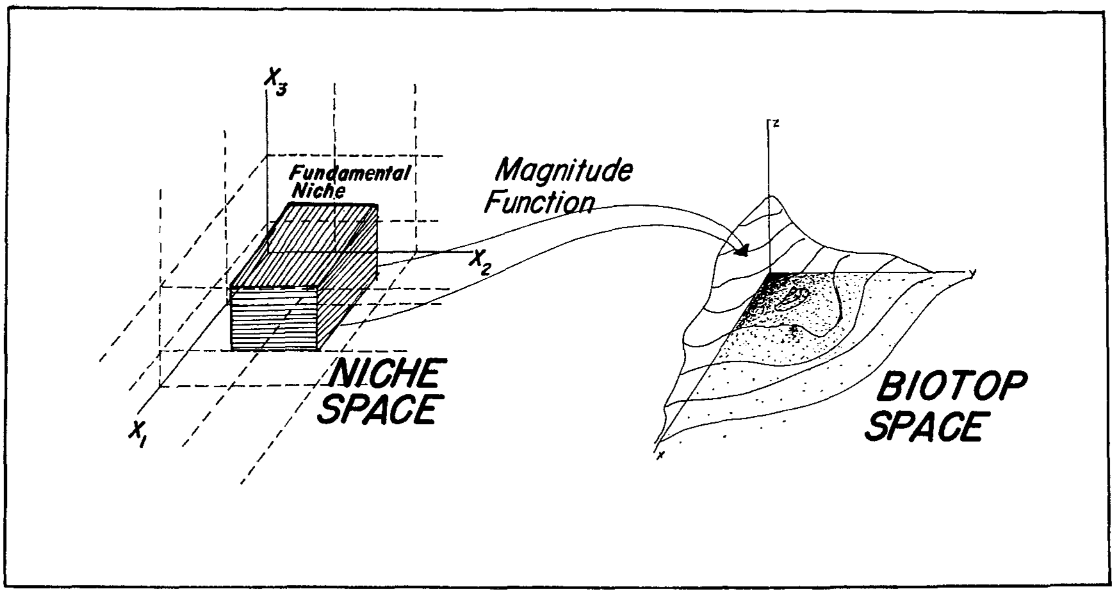

FIG. 1. The magnitude function, $f(x_1, x_2, x_3)$ determines the density, (z), of settlement in any area, dB, of the biotope.

in the biotope space, corresponding to some phytophysiological limit in the niche space.

This assumes that all vectors in the fundamental niche describe conditions equally favorable to settlement. Or in other words, there are no gradients in the spatial distribution. If this were true, Venn diagrams would be useful to describe the preceding concepts. However, gradients are very important. In general the function Φ defines a mapping that determines the density of settlement in the biotope space and is called the magnitude function. This function determines the density, but not the pattern of settlement in any area dB of the biotope. Later it will be shown that quite different configurations result from alternative assumptions about the magnitude function and how it varies from place to place.

A simpler but less general model of colonization may be illustrated by a Thünen-type model of settlement density, where the gradient function is monotone decreasing, but not necessarily linear (Fig. 2).

In this case the magnitude function defines a single-valued mapping from a one-dimensional niche space to a magnitude of occurrence which varies inversely with distance. Settlement density is a function of a single variable, x, making density decrease with distance from an optimum spot in the environment. The total number of inhabitants occupying the biotope within an area of radius r of the optimum point is

$$(1) \qquad n_r = 2\pi \int_0^r kx\,f(x)\,dx$$

where k is a constant, depending on the units of measurements used. The total settlement in the area is found by integrating over all x such that $f(x) > 0$. At the point a distance x_0 from O such that $f(x_0) = 0$, the outer limit of settlement is reached. The number of inhabitants in the annulus $(x, x + \Delta x_i)$ is given by the volume of the cylindrical shell,

$$(2) \qquad n_i = 2\pi kx\,f(x)\,\Delta x_i.$$

If a change in limits on the fundamental niche occurs, then the domain of $f(x)$ will change. Under colonization alone, with no increase in numbers, the biotope will cover a larger area with the same number of settlements. The settlers on the frontier come from some place in the biotope, thereby lowering density, in order to expand the frontier. If conditions become more restrictive, $f(x)$ may steepen, and frontier settlement will be abandoned. It is characteristic that frontier zones

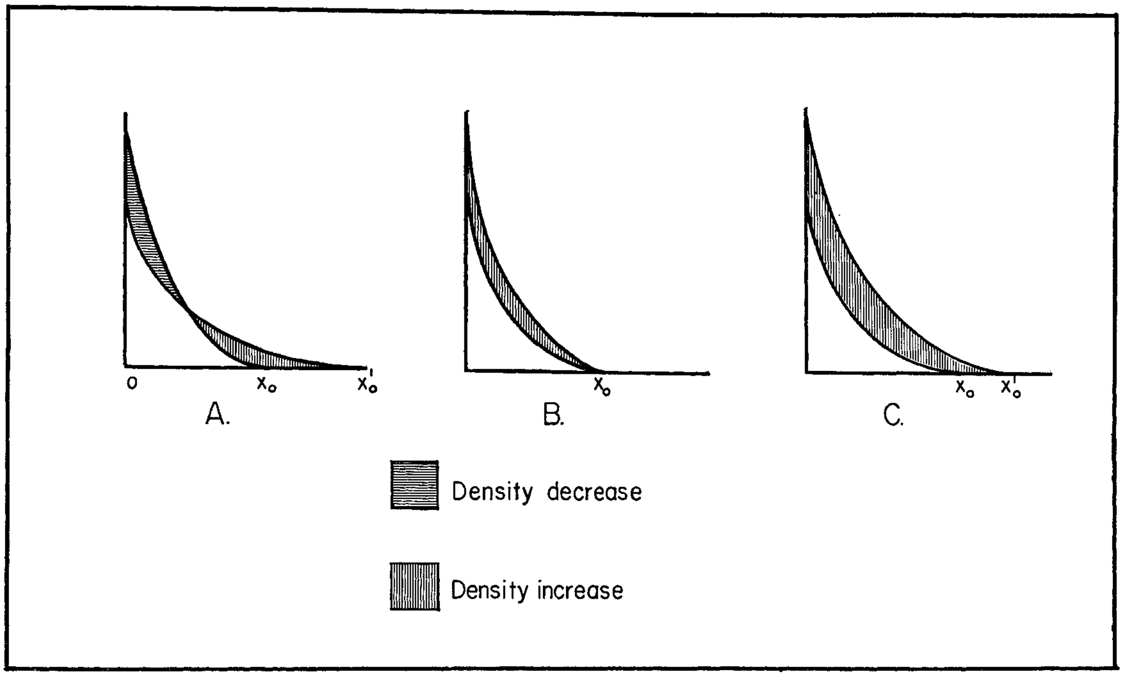

FIG. 2. The relationship between colonization and spread when the niche is one-dimensional: A) Colonization alone: the magnitude function shifts outwards from x_o to x'_o; B) spread alone: the magnitude function shifts upwards; C) colonization and spread: the function shifts upwards, and outwards from x_o to x'_o.

are the most unstable.[14] Colonization is thus a function of the size of the fundamental niche. As this expands, settlement may spread outward, and if it contracts, frontier settlement is abandoned.

Spread

Spread as mentioned above, is one of two phases of a process producing diffusion of settlements. This process is characterized by a general filling up of the biotope by settlers. In the process of colonization, the parameters in the niche space change, permitting migration into unoccupied territory. In the process of spread, however, the magnitude function is shifting upwards, but not outwards. It is a process which produces an increase in the density, whereas colonization in effect decreases density by increasing the area of the biotope.

In ecological models of spread, the nonspatial process is usually population growth. The spatial process is diffusion, but most critically a compound diffusion resulting from successive offspring replacing themselves and adding new individuals to the population. The 1957 Neyman and Scott model of population distribution embodies many of these characteristics.[15]

Bylund has produced several models of the processes that have been discussed here, resembling the simulation models for innovation diffusion developed by Hägerstrand.[16] Bylund gave evidence of the importance of Spread in explaining the distribution of settlement. He noted that in the Lappish lands there have not been waves of settlement but:[17]

> . . . rather only a handful of persons, who appeared as pioneers in the wilderness. Afterwards the

[14] O. E. Baker, "Government Research in Aid of Settlers and Farmers in the Northern Great Plains of the United States," in W. L. G. Joerg (Ed.), *Pioneer Settlement* (New York: American Geographical Society, Special Publication No. 14, 1932), p. 77.

[15] J. Neyman and E. Scott, "On a Mathematical Theory of Populations Conceived as a Conglomeration of Clusters," *Cold Spring Harbor Symposia on Quantitative Biology,* Vol. 22 (1957), pp. 109–20.

[16] Bylund, *op. cit.,* footnote 8, pp. 225–31.

[17] Bylund, *op. cit.,* footnote 8, pp. 225–26.

following colonization has been carried out by sons of the first pioneers and then their sons, and etc. I have called this development clone-colonization (clone from the Greek word clon-branch) i.e. a development from one genealogical tree. Before 1870 about three fourths of the settlements in Pite lappmark were results of the clone-colonization, or an inner colonization, not the result of immigration.

A nonsimulation approach to the process was provided in 1939 by the statistician, Neyman, who was concerned with predicting the spread of larvae from a small area around their birth place.[18] His work may be rephrased in terms of human settlement.

Neyman's model is an extension of the Poisson probability law for settlements dispersing from randomly distributed groups. The problem is to find the probable number of settlers in an area of the biotope at some specified future time.

If there are N original settlers in a total area M of the biotope; let $\alpha = M/N$. The average number of individuals surviving from a settlement at a given time is denoted, β. Also, it is assumed that there is an area from which the settlers found in unit area w of M must have come, Γ.

The Neyman Type A distribution has two parameters:

$$m_1 = \alpha\Gamma$$
$$m_2 = \beta/\Gamma$$

m_1 is the mean number of clusters per unit area and m_2 is the mean number of units per cluster. The probability generating function of the random variable x is shown by Neyman to be:

$$(3) \quad P(x = n) = \frac{e^{-m_1} m_2^n}{n!} \frac{d^n}{d^n u} e^{m_1 e^{u - m_2}} \Bigg|_{u=0}.$$

A computational form was presented by Neyman, and also by Beall.[19] The probability that the number of settlements, x, will be $n + 1$ is:

$$(4) \quad P(x = n + 1) = \frac{m_1 m_2 e^{-m_2}}{n + 1} \sum_{k=0}^{n} \frac{m_2^k}{k!} P(x = n - k)$$

which starts with

$$(5) \qquad P(x = 0) = e^{-m_1(1 - e^{-m_2})}.$$

The assumption of Bylund's and Neyman's model is that successive generations show a limited spread away from birthplace. This is such a widely used postulate in migration models in all sciences that it probably does not warrant further elaboration here.[20]

Despite Bylund's observation, it is clear that a principal determinant of increasing density in many frontier zones is not spread but continued migration to the frontier from outside the area. More important geographically, there is no reason to expect this new immigration to cluster around the settlements of the pioneers. In fact, it seems likely that new settlement would be somewhat repelled by the earlier settlement, under conditions of contiguous landholdings of approximately equal size typical of most homesteading in the United States. The result would eventually be greater regularity in the spacing of farmsteads, rather than clustering.

In short, the two processes of colonization and spread may occur not only in the same place, but also at the same time. The morphology of settlement affected by colonization alone possibly produces a regular spacing whereas the pattern arising from reproduction alone tends to clustering, especially after several generations.[21]

Competition

No matter whether density of some small area of the biotope increases as a result of colonization or spread, there occurs an upper limit to the area of the biotope, which checks growth. Ecologists refer to these types of

[18] J. Neyman, "On a New Class of 'Contagious' Distributions Applicable in Entomology and Bacteriology," *Annals of Mathematical Statistics*, Vol. 10 (1939), pp. 35–37.

[19] G. Beall, "Fit and Significance of Contagious Distributions When Applied to Observations on Larval Insects," *Ecology*, Vol. 21 (1940), pp. 460–74.

[20] For example, see Neyman and Scott, *op. cit.*, footnote 15, p. 109; R. L. Morrill, "The Distribution of Migration Distances," *Papers of the Regional Science Association*, Vol. 11 (1963), pp. 75–84.

[21] Hutchinson terms this "Reproductive Pattern." See G. E. Hutchinson, "The Concept of Pattern in Ecology," *Proceedings of the Academy of Natural Sciences of Philadelphia*, Vol. 105 (1953), pp. 1–12.

checks on population as density-dependent conditions.[22] In animal populations, density-dependent conditions are those of exhaustion of the available food supply. The counterpart in rural settlement is that there is a lower limit on the size of farm that can be operated economically. Increase in prices may bring smaller farms into production (change in volume of the niche space), but even with such a condition, the biotope is of finite size and so are the farms, and the process is inevitably checked.

The process of competition is a struggle between settlements to hold their domains intact and to increase their holdings. Larger settlements absorb smaller ones, just as large trees get greater nourishment by blocking sunlight from smaller trees near them.[23] Trade centers do the same thing. There exists a finite amount of purchasing potential spread over the hinterlands. Towns must compete with one another for customers, extending their trading areas as far as possible. Those centers with a disadvantageous position on the map are squeezed out.

When the biotope is not completely covered with farms, new settlers, whether they are the result of colonization or spread, may carve out their own portion of the biotope in a manner which does not affect the size or position of those already present. Under low density conditions there is no need for competition. It is only under high density conditions that the competition for space comes into being and the process will become important only when farmers try to expand their holdings. The recent agricultural history of the United States has been characterized by these conditions—a small decrease of land in farms, a marked increase in average size of farm, and a decline in the number of farms.[24]

When a farmer acquires more land he wishes to keep his holdings compact, to avoid unnecessary travel, thereby lowering costs.[25] He wishes to be located closest to his place of work. The extremum problem involved may be phrased as "locate farmers closest to their farms."[26] The optimum pattern of farmsteads is the hexagonal lattice, when all farms are of the same size. Under equal-size holdings, locating farmers closest to the land is the same as locating them farthest from each other. However, locating farmers closest to their individual farms may be accomplished without the farms being of the same size, hence the settlement pattern need not be regular. The nature of the size distribution of farms thus enters the problem. If the magnitude function governing settlement density is a constant over a large area of the biotope, then there is no *a priori* reason for assuming that farms will be of different sizes. Any such differences may be left as chance factors in the present model and might involve entrepreneurial skill or availability of capital to certain farmers. A constant magnitude function implies that any area of the biotope will be as capable of supporting a given population density as any other. To the extent that all settlers are equally competitive the farms will be of equal size. Equilibrium is reached when no settler is able to absorb another settler's land.

COMPETITION AND CENTRAL PLACE THEORY

Although stochastic models of settlement spread were described briefly above, no probabilistic formulation of the entire process was attempted. Now a probabilistic interpretation of competition will be introduced. It will be shown that the assumption of a constant magnitude function is not a sufficient condition for the central place theory assumption of uniform base population, under a reasonable set of assumptions about settlement density.

The exact nature of the uniformity of the underlying population assumed in central place theory seems to be a matter of question.

[22] H. G. Andrewartha and L. C. Birch, *The Distribution and Abundance of Animals* (Chicago: University of Chicago Press, 1954), p. 16.

[23] This illustration of competition is given in: M. Maruyama, "The Second Cybernetics: Deviation-Amplifying Mutual Causal Processes," *American Scientist*, Vol. 51 (1963), p. 167.

[24] U. S. Department of Commerce, Bureau of the Census, *Historical Statistics of the United States* (Washington, D. C.: Government Printing Office, 1960), pp. 278–80.

[25] Data on the cost of dispersed operations is given in M. Chisholm, *Rural Settlement and Land Use* (London: Hutchinson University Library, 1962), p. 55.

[26] This is the same as Bunge's extremum problem "Locate points in an area as near to the area as possible." See Bunge, *op. cit.*, footnote 4, p. 225.

It is not clear what geometrical properties Christaller wanted the farms to assume. Lösch assumed that the uniformity was hexagonal and constructs a proof to show that this arrangement yields the demand cone of largest volume.[27] Beckmann, however, stated:[28]

> The very simplest model of a city hierarchy would run along lines somewhat like the following. There is a basic layer of rural population settled at a uniform density, or alternatively there is a random scattering of the smallest communities with an approximately uniform areal density. The first layer of cities superimposed on this basis consists of centers performing the most elementary production and distribution function.

In a probability sense, what Beckmann described corresponds to a Poisson probability distribution, which would fulfill the condition of uniform areal density. Assuming that the cones have a sufficiently large base, then the introduction of hexagonal spacing of base population into Lösch's argument would be unnecessary, granting slight deviations.

In geometrical probability, the problem of distribution of random nonoverlapping circles all of the same size on a plane, was discussed by Kendall and Moran.[29] Assuming that farms are compact, the problem has relevance in settlement theory.

If all circles have diameter d, and one circle has its center at P, then there cannot exist another circle centered closer than d to P. The likelihood of a circle having its center in the infinitesimal area ds, where this area is a distance R from P, $(R > d)$, is not a function of ds alone. Also, it is not a monotonic function of R. Kendall and Moran point out that under close packing the probability of a circle center falling in ds at a distance R from P, is a strongly peaked function of R. Apparently no solution to this problem in geometrical probability has been found.

A contribution by Pielou is of great value in understanding the problem.[30] Pielou constructed an analog model of plant distribution by randomly selecting points on a large square, and then drawing circles around them such that no circles were allowed to overlap. She did this by deleting all subsequently chosen points falling within or too close to previously determined circles. The size of circles was allowed to vary in her experiments. In model A, circle radius ranged between .2 and .7 units, in model B the range was .1 to 1.0 units and in model C, the range was .1 to 2.0 units. It was found that:[31]

> . . . the models of set A remained regular even though the density was increased until the sheet was almost filled. With set B, however, in which the permitted range of circle size was larger, the variance-mean ratio increased with increasing density until for high densities it did not differ significantly from unity, suggesting that these populations were random. In the set C, that with the largest range of circle sizes, apparently aggregated population were finally obtained, and the degree of aggregation increases with increasing density.

Assuming that the settlement units are highly compact, and rather closely packed, Pielou's findings shed much light on the problem of regular or random spacing. A large range in farm sizes produces a clustered settlement form. If the range in farm sizes is moderate, a random distribution will occur, whereas a regular distribution results when variance is small.

Suppose that size-of-farm in acres is a random variable, x, with a probability density function,

$$(6) \qquad f(x) = \frac{ck^c}{x^{c+1}} \qquad \begin{array}{r} k, c > 0 \\ x > k \end{array}$$
$$= 0 \text{ when } x < k.$$

This is a Pareto distribution and is linear in the logs with a slope of $-(c+1)$.[32] Such a curve is one statement of the rank-size distribution and its wide applicability seems to make it a logical one to apply to farm sizes.

The expectation of the random variable x is

$$(7) \qquad E(x) = \int_k^\infty \frac{ck^c\,dx}{x^c} = \frac{ck}{c-1}$$

[27] Lösch, *op. cit.*, footnote 7, pp. 105–23.

[28] M. Beckmann, "City Hierarchies and the Distribution of City Sizes," *Economic Development and Cultural Change*, Vol. 6 (1958), p. 243.

[29] M. G. Kendall and P. A. P. Moran, *Geometrical Probability* (London: Griffins, 1963), pp. 46–47.

[30] E. C. Pielou, "A Single Mechanism to Account for Regular, Random and Aggregated Populations," *Journal of Ecology*, Vol. 48 (1960), pp. 575–84.

[31] Pielou, *op. cit.*, footnote 30, p. 577.

[32] H. D. Brunk, *Introduction to Mathematical Statistics* (New York: Blaisdell, 1965), p. 55.

and is defined for any $c > 1$. The variance is defined as

$$(8) \qquad \int_k^\infty \frac{ck^c\,dx}{x^{c-1}} - \frac{c^2k^2}{(c-1)^2} = \frac{ck^2}{c-2} - \frac{c^2k^2}{(c-1)^2} \text{ when } c > 2.$$

Taking

$$\lim_{c\to\infty} \frac{k^2}{c^2 - 4c + 5 - 2/c} = 0.$$

When *c* becomes very large, the variance of the distribution approaches zero. In this case, the rank-size plot would slope downward very rapidly. In the limit, this distribution has a one-point distribution, and it is shown by Fisz that zero variance is a necessary and sufficient condition for a probability density function to be one-point.[33] Thus, a necessary and sufficient condition for the farm size (rank-size) distribution to approach the central place model, and *vice versa*, is that the size distribution have a rank-size plot with a slope approaching infinity—a single size of farm.

The precise difference between the uniform distribution of settlements assumed in central place theory and uniform areal settlement density is explained in terms of the limiting forms of probability distributions describing the two cases.[34] Dacey showed that the distribution of the standardized *j*th order distance in a population of *n* randomly distributed individuals converges asymptotically to a two dimensional areal uniform distribution. For a given *j* the correspondence becomes greater as *n* becomes large, or, as density increases.

However, Dacey noted that probably the greatest interest lies in the case where density is low. The correspondence between the discrete random and areal uniform distribution is smallest in a spatial distribution such as rural settlement. This means that in a settlement pattern affected by competition, the low areal density requires that in order for settlements to have a uniform distribution their areal variance must be smaller than that which would result simply from uniform areal density—the pattern must be more regular, such as a discrete uniform distribution, an extreme form of which is the hexagonal lattice. Whether or not actual settlement distributions are adequately described by a model assuming only a constant magnitude function or are more regular than such a model allows, is a matter for empirical investigation. From the arguments presented above, it is clear that there are conditions under which a highly regular settlement pattern is not to be expected. Specifically, when settlement is affected mainly by colonization and spread, the necessary conditions for regularity do not exist. Either when density is insufficient to cause competition for space or when clone colonization is dominant, the theory states that a regular structure is unlikely. The morphological laws of diffusion, which predict the existence of settlement clusters, are of interest in the latter case.

MATHEMATICAL MODELS OF SETTLEMENT PROCESS

It is desirable to reformulate the settlement models in a more tractable mathematical form than was used in the general explication of the theory. This will permit testing the generalizations against actual occurrences to determine their usefulness in predicting characteristics of settlement location.

If it is assumed that the biotope is homogeneous in all respects and the differences in the size of settlements (farms) owe entirely to a set of variables whose net effect is random, then the magnitude function

$$\Phi(X_i) = m$$

a constant density for every

$$(x_{i1}, x_{i2}, \ldots x_{in}) = (X_i) \text{ in } N.$$

Further assume that no process of competition exists that conditions the spacing of settlements, such that the size and position of one farm in no way affects another farm's location, except that it is unlikely that they are extremely close together. These assumptions are sufficient to specify a Poisson probability process. Specifically, from the

[33] M. Fisz, *Probabililty Theory and Mathematical Statistics* (New York: John Wiley and Sons, 1963), p. 129.

[34] M. F. Dacey, "Two Dimensional Random Point Patterns: A Review and an Interpretation," *Papers of the Regional Science Association,* Vol. 13 (1964), p. 54.

following three assumptions, the Poisson law may be derived:[35]

1) The probability of a settlement occurring in an area, a, is a function only of the size of a, not its location in the study area; 2) the number of settlements in any small part of the study area is independent of the number falling in any other area; 3) the probability of more than one settlement occurring in a, approaches zero as a approaches zero, faster than a does.

The result is the probability density function describing the likelihood of finding exactly x settlements in a,

$$f(x) = \frac{m^x e^{-m}}{x!} \qquad x = 0, 1, 2, \ldots \tag{9}$$
$$= 0 \text{ elsewhere.}$$

Suppose that the settlement process is not characterized by such independence of locations, but rather by areas of the biotope where greater settlement density is likely, owing to spread or favorable physical features and by areas of low density, as when physical restraints such as poor soils or hilliness, are found. In this case it might be assumed that m has a gamma-type distribution,

$$g(m) = \frac{1}{\Gamma(\alpha)\beta^\alpha}(m^{\alpha-1}e^{-m/\beta}). \qquad 0 < \beta, m \tag{10}$$
$$= 0 \text{ elsewhere.} \qquad \alpha = 1, 2, 3, \ldots$$

The function $f(x)$ is now redefined as a conditional probability density function

$$f(x \mid m) = \frac{m^x e^{-m}}{x!} \qquad x = 0, 1, 2, \ldots$$

with the probability distribution of x depending on the *a priori* distribution of m. Then the joint probability density function becomes,

$$h(x, m) = \frac{1}{\Gamma(\alpha)\beta^\alpha x!}(m^{x+\alpha-1}e^{-m(1+1/\beta)}), \qquad x = 0, 1, 2, \ldots \quad m, \beta > 0$$
$$= 0 \text{ elsewhere.} \qquad \alpha = 1, 2, 3, \ldots$$

By integrating out the parameter, m, the negative binomial distribution is obtained.[36]

[35] Fisz, *op. cit.*, footnote 33, pp. 276–79.

[36] M. Greenwood and G. Yule, "An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents," *Journal of the Royal Statistical Society*, Vol. 3 (1920), pp. 255–79.

Let $w = m(\beta + 1/\beta)$. Then the new marginal probability density function of x is

$$h(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha x!}\left(\frac{\beta}{\beta+1}\right)^{x+\alpha} \int_0^\infty w^{x+\alpha-1}e^{-w}\,dw,$$

$$= \binom{x+\alpha-1}{\alpha-1}\frac{\beta^x}{(\beta+1)^{x+\alpha}} \tag{11}$$

which is a negative binomial distribution with

$$0 < t = \beta/1 + \beta < 1.$$
$$0 < s = 1/1 + \beta < 1.$$
$$\alpha = r = 1, 2, 3, \ldots$$

Rewriting equation (11) yields the negative binomial density:

$$h(x) = \binom{x+r-1}{r-1} s^r t^x \qquad x = 0, 1, 2, 3, \ldots$$

The smaller the value of r, the greater the amount of clustering.[37] The case of regular spacing in settlement may also be considered as a Poisson process, but with a bias toward evenness. Depending on the strength of competition, there will be a greater or lesser amount of dependence between settlement locations. Define a parameter called "evenness bias," p, such that $0 \leqslant p \leqslant 1$, with large values of p indicating evenness, and its complement $q = 1 - p$, a randomness "bias." In this case, m, is a constant, but it is desired to make the areal variance small. Dacey's regular Poisson modification is appropriate.[38] The probability density function is

$$p(x) = qe^{-\gamma} \qquad x = 0 \tag{12}$$
$$= \frac{q\gamma^x e^{-\gamma}}{x!} + \frac{p\gamma^{(x-1)}e^{-\gamma}}{(x-1)!}$$
$$= 0 \text{ elsewhere.} \qquad x = 1, 2, \ldots$$

The parameters are $\gamma > 0$; $0 \leqslant p \leqslant 1$; $q = 1 - p$; where $p = (\bar{x} - \sigma_x{}^2)$, and $\gamma = \bar{x} - p$.

[37] This parameter has been used by Rogers to interpret the degree of clustering of similar land use types within the city. See A. Rogers, "A Stochastic Analysis of the Spatial Clustering of Retail Establishments," *Journal of the American Statistical Association*, Vol. 60 (1965), pp. 1094–1103.

[38] M. F. Dacey, "Modified Poisson Probability Law for Point Pattern More Regular than Random," *Annals*, Association of American Geographers, Vol. 54 (1964), pp. 559–65.

If there is no competition, $p = 0$ and the series reduces to (9). Consider a square settlement grid each cell of which may contain 0, 1, 2, . . . settlements. If $p = 1$, the probability that a square will have no settlement is zero, which is not true of the ordinary Poisson model. The probability is concentrated among low values of x.

An important distinction between the Poisson, negative binomial, and regular Poisson distributions is the differences in their variance-mean ratios, defined as variance divided by the mean. In practice, an actual settlement distribution is enumerated by laying a grid of quadrats over the spatial distribution and counting the number of settlements in each quadrat. The frequency distribution of these quadrat densities has a definite mean and variance. When the mean is larger than the variance, the distribution is termed regular; when the variance is larger than the mean it is termed clustered. The term random is applied to the case when variance and mean are equal, *i.e.*, the variance-mean ratio is one.

The means, variances, and variance-mean ratios of the probability models employed here are shown in Table 1.

The regular Poisson distribution will be fit to settlement data when the empirical variance-mean ratio is smaller than or equal to one, and the negative binomial will be fit when the ratio is greater than or equal to one, whereas the Poisson distribution will be fit to all settlement maps. The variance-mean ratio thus provides an indicator of what probability distribution applies. It also has a known sampling distribution, and the significance of its departure from unity may be tested.[39]

TABLE 1.—MOMENTS OF THE PROBABILITY MODELS*

Distribution	Mean	Variance	Variance-mean ratio
Poisson	m	m	$m/m = 1$
Regular Poisson	$m + p$	$m + p - p^2$	$\frac{m + p - p^2}{m + p} < 1$, when $p > 0$.
Negative binomial	rt/s	rt/s^2	$1/s > 1$, when $0 < s < 1$.

* See text for explanation of symbols.
Source: compiled by author.

[39] P. Greig-Smith, *Quantitative Plant Ecology* (London: Butterworths, 1964), p. 63.

Empirical Tests

Settlement data from several areas of eastern Iowa at three times were collected and examined for evidence of the morphology suggested by theory (Fig. 3). The sample consisted of tracts in six counties: Bremer, Cedar, Johnson, Jones, Henry, and Tama. The counties were selected because of data availability. The frontier had passed each of these counties by the time the earliest settlement data could be obtained.[40] For this reason there could be no test of the theory regarding colonization as a process of expansion of the frontier.

A further reason for selecting these study areas was that little topographic influence is present. Thus it can be interpreted that if settlements are clustered the cause should not be environmental restraints, but instead a contagious process of settlement growth.

The study areas are tracts ranging in area from 104 to 200 square miles. Each study area was divided into two approximately equal areas for purposes of data analysis. No incorporated towns are located within the areas. A total of thirty-six spatial distributions were analyzed, made up of the dual tracts in six counties at three time periods.

The settlement map was the sole source of data used. Each settlement (farmhouse) was recorded in cartesian coordinates on a grid consisting of a 100×100 square lattice superimposed on each square mile.[41] For the first two time periods the settlement maps were those of county commercial atlases. For the most recent data, county highway maps published by the Iowa Highway Commission were used. Each of these sources distinguished between farm and non-farm residences, and noted abandoned dwellings. Only occupied

[40] This judgment is based on a map in the 1890 Census of Population in which settlement isochrones showing the position of the frontier at various times were drawn. The limit of settled territory is defined as two persons per square mile. Under this definition, the frontier passed all of the study areas except Tama and Bremer counties by 1850, and passed these two shortly afterwards. See: U. S. Bureau of Census, *Report on Population of the United States: Eleventh Census, 1890, Part 1* (Washington, D. C.: Government Printing Office, 1895).

[41] I am indebted to my wife, Beverly, who recorded the settlement locations and punched the data on cards.

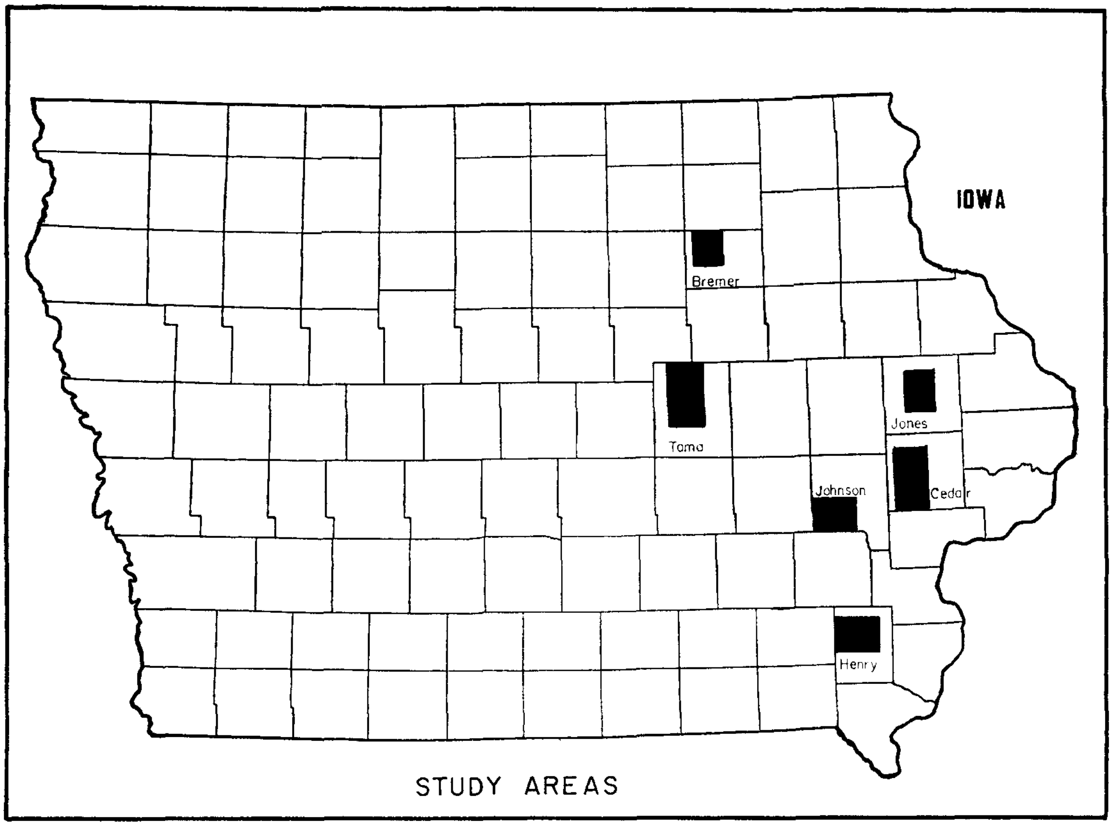

Figure 3.

farm residences were recorded, save for Cedar County in 1940, where all occupied residences were enumerated for sake of comparison. Field checking revealed that the recent maps were approximately ninety percent accurate. The principal errors involve failure to note farm abandonment. The density of farms is being somewhat overestimated, at least for the most recent maps.

A computer program was written which read the cards describing each location, enumerated the distribution with a square grid of quadrats and fit the Poisson, negative binomial, and regular Poisson distributions. Then, it computed the Kolmogorov-Smirnov one sample goodness-of-fit test for each of the three pairs of observed-expected distributions.[42] Finally the variance-mean ratio was computed and a t-test of its significance was carried out. The procedure was repeated for any number of quadrat sizes.[43]

Since two kinds of tests were made, the interpretation of the results is unclear in some cases, as when the variance-mean ratio indicates a configuration different from that suggested by the fits of the probability distributions. The reason for computing two tests is to obtain more information, since the two approaches have rather different properties and impart some different information.

The technique described above, sometimes referred to as cell count analysis, is highly sensitive to quadrat (cell) size. Considerable doubt is cast upon the findings of researchers who have used a single cell size,

[42] S. Siegel, *Non-Parametric Statistics* (New York: McGraw-Hill, 1956), pp. 47–52.

[43] The program was coded in FORTRAN IV for the IBM 7044 at the University of Iowa. A print-out of the program, entitled MAPPI, may be obtained from the author. Computation time was made available through the Graduate College of the University of Iowa.

especially when the size had no theoretical relation to the problem being studied.[44] Also, it is easily shown that different frequency distributions may in certain cases produce identical variance-mean ratios.[45] To compound the problem, different spatial distributions produce identical probability series.[46]

To combat these sources of error, the probability models provide an alternative to the variance-mean ratio. Also, the probability models were fit at different quadrat sizes. This eliminates quadrat-size bias to a considerable extent, and produces evidence on another important fact—the scale at which clustering or regularity takes place, *vis-à-vis* quadrat size. One quadrat size was selected to correspond approximately to the average size of the farm. The quadrats were 1/16, 1/4, and one square mile in area, or 40, 160, and 640 acres in size. The 160-acre quadrat is roughly equal to average farm size.

Some geographers have given the erroneous impression that there exists such a thing as a "truly random" spatial distribution, supposedly meaning a spatial distribution that best fits the Poisson series at any quadrat size.[47] The reduction in variance with the increase in quadrat size, or its equivalent, an increase in density at a constant quadrat size, has long been known to plant ecologists.[48] In general, the larger the quadrat size, the greater the regularity, hence the lower the variance of the frequency distribution of cell counts. Distributions are random, regular, or clustered only with respect to a given quadrat size and size of study area in the case of cell count analysis, and with respect to study area size in near-neighbor analysis.[49] All subsequent statements about the randomness, regularity, or clustering of rural settlement distributions should be interpreted with these limitations in mind.

From the theory of settlement processes, three hypotheses were formulated to be tested in the study areas reported above.

1) The process of competition among fixed points lowers density. It is hypothesized that in a distribution where density has been reasonably high but where there has been a subsequent decline in density, the distribution will tend towards a more regular configuration, as the average size of farms increase.
2) Greater occurrence of settlement clustering indicates an increase in density over a state in which such clustering was not present. Recall, however, that increasing density can also be produced by continued migration from the outside. The occurrence of an increasing degree of clustering is a sufficient but not a necessary condition for increasing density. Hence the form of the hypothesis: If clustering increases in a spatial distribution, then density should have increased.
3) In an area undergoing farm abandonment and an increase in farm size, but that once supported a larger farming population, in the long run the process of competition should be most in evidence. The pattern should become increasingly orderly over time, assuming that most land remains in farms. This hypothesis is testable in Iowa, since the farm economy did experience these adjustments during the period studied.

Hypotheses 1 and 3 differ since the first is concerned with short-term changes in density as related to settlement aggregation, whereas the latter is concerned with long-term trends as affected by economic and other conditions.

---

[44] This is one of the deficiencies in Harvey's study of spatial diffusion using cell-count analysis. See D. Harvey, "Geographic Processes and the Analysis of Point Patterns: Testing Models of Diffusion by Quadrat Sampling," *Transactions of the Institute of British Geographers,* Vol. 40 (1966), pp. 81–95.

[45] K. A. Kershaw, "The Detection of Pattern and Association," *Journal of Ecology,* Vol. 48 (1960), pp. 233–42.

[46] M. F. Dacey, "A County-Seat Model for the Areal Pattern of an Urban System," *Geographical Review,* Vol. 56 (1966), p. 529.

[47] See G. Olsson, "Central Place Systems, Spatial Interaction, and Stochastic Processes," *Papers of the Regional Science Association,* Vol. 18 (1967), pp. 13–45; H. McConnell, "Quadrat Methods in Map Analysis," *Discussion Paper Series* (Iowa City: Department of Geography, University of Iowa, No. 3, 1966).

[48] K. A. Kershaw, *Quantitative and Dynamic Ecology* (New York: American Elsevier, 1964), p. 106.

[49] J. C. Hudson and P. M. Fowler, "The Concept of Pattern in Geography," *Discussion Paper Series* (Iowa City: Department of Geography, University of Iowa, No. 1, 1966).

TABLE 2.—VARIANCE-MEAN RATIOS AND DENSITIES

Time period	Quadrat size and density	Study area: Bremer West	Bremer East	Cedar South	Cedar North	Henry West	Henry East
	Year	1875		1872		1870	
	Density	279	227	170	187	408	363
I	1/16 sq. mi.	.976	.879*	1.048	.939	1.017	1.970*
	1/4 sq. mi.	.945	.789*	1.179	.914	.997	1.196*
	1 sq. mi.	1.143	1.165	1.417*	.915	1.105	1.409*
	Year	1917		1901		1917	
	Density	291	237	173	175	322	312
II	1/16 sq. mi.	1.040	.999	.991	.942	.857*	.991
	1/4 sq. mi.	1.020	.919	.951	.815	.760*	.682*
	1 sq. mi.	1.432*	.913	1.082	.703*	.779	.604*
	Year	1960		1940		1960	
	Density	264	264	469	380	267	290
III	1/16 sq. mi.	.837*	.815*	1.923*	1.240*	.850*	.858
	1/4 sq. mi.	.788*	.591*	3.136*	1.335*	.714*	.762*
	1 sq. mi.	.813	.412*	4.039*	1.034*	.796	.554*

Time period	Quadrat size and density	Johnson South	Johnson North	Jones South	Jones North	Tama South	Tama North
	Year	1870		1893		1875	
	Density	317	229	222	270	291	178
I	1/16 sq. mi.	.946	1.012	1.010	.978	1.024	1.018
	1/4 sq. mi.	.899	1.073	1.120	.869	.999	1.026
	1 sq. mi.	1.117	.915	1.323	.751	1.348*	1.169
	Year	1917		1915		1916	
	Density	363	262	217	242	220	287
II	1/16 sq. mi.	1.005	.987	.986	1.013	.895*	.877*
	1/4 sq. mi.	.889	.880	.870	.990	.596*	.679*
	1 sq. mi.	.852	.885	.904	.960	.693*	.727*
	Year	1960		1960		1960	
	Density	358	249	175	245	301	256
III	1/16 sq. mi.	.992	.881*	.899*	.907*	.828*	.866*
	1/4 sq. mi.	.838*	1.023	.893	.709*	.494*	.524*
	1 sq. mi.	.940	.765	1.126	.628*	.331*	.660*

* Indicates variance-mean ratio is significant from unity with a t-test, one tail, for $\alpha = .05$.
(In Cedar County, 1940, all residences were recorded, including several hamlets. The result is a much higher degree of clustering.)
Source: compiled by author.

Hypothesis 1 proposed that the process of competition and its concomitant effect of lowering density (farm abandonment), would produce greater regularity in the distribution. This is definitely the case in Henry County, for both time intervals (see Tables 2 and 3). With two exceptions, variance-mean ratios steadily declined as did the density, between 1872 and 1960. The same is true for southern Jones County, save for a tendency toward avoidance of some areas in 1960. In the south unit of Bremer County between 1917 and 1960 the same occurred, and also in the north unit of the Cedar County study area, between 1872 and 1901. In northern Jones County between 1893 and 1915, the density decreased and clustering increased, but the variance-mean ratios are not significant from

TABLE 3.—SIGNIFICANT FITS OF THE PROBABILITY MODELS** (Best fits are underlined)

Time period	Study area	Quadrat size 1/16 sq. mi.	1/4 sq. mi.	1 sq. mi.
	Bremer west	P RP	P RP	P NB
I	Bremer east	P RP	P RP	NB
	Cedar South	P NB	P NB	NB
	Cedar north	P RP	P RP	P RP
	Henry west	P NB	P RP	None
	Henry east	P NB	P NB	NB
	Johnson south	P RP	P RP	P NB
	Johnson north	P NB	P NB	P RP
	Jones south	P NB	P NB	P NB
	Jones north	P RP	P RP	P RP
	Tama south	P NB	P RP	NB
	Tama north	P NB	P NB	P NB
	Bremer west	P NB	P NB	None
II	Bremer east	P RP	P RP	P RP
	Cedar south	P RP	P RP	P NB
	Cedar north	P RP	P RP	*
	Henry west	P RP	P RP	*
	Henry east	P RP	P RP	RP
	Johnson south	P NB	P RP	P RP
II	Johnson north	P RP	P RP	P RP
	Jones south	P RP	P RP	P RP
	Jones north	P NB	P RP	P RP
	Tama south	P RP	RP	P RP
	Tama north	P RP	RP	P RP
	Bremer west	P RP	P RP	RP
III	Bremer east	P RP	RP	*
	Cedar south	P NB	P NB	P NB
	Cedar north	P NB	P NB	P NB
	Henry west	P RP	P RP	P RP
	Henry east	P RP	P RP	RP
	Johnson south	P RP	P RP	P RP
	Johnson north	P RP	P NB	P RP
	Jones south	P RP	P RP	P NB
	Jones north	P RP	P RP	*
	Tama south	P RP	P RP	*
	Tama north	P RP	RP	*

P: Poisson distribution; NB: Negative Binomial; RP: Regular Poisson.

* indicates inequality violated for regular Poisson distribution (see footnote 50).

** significant fits, under the Kolmogorov-Smirnov test, for $\alpha = .05$.

Source: compiled by author.

unity. In Tama County density decreases occurred in the south unit between 1875 and 1916, and in the north unit between 1916 and 1960. In both cases, the distribution tends toward greater regularity with farm abandonment. In Johnson County-north between 1917 and 1960, two of the three quadrat sizes indicate increased regularity, although the density declines by less than five percent. In the other cases, density decreased by less than three percent which is probably beyond the margin of error in data collection.

Thus, in the eleven cases where the hypothesis, "decreased density accompanies increased regularity" could reasonably be tested, the predicted result occurred in ten and was probably not borne out in one.

Hypothesis 2 proposed that an increase in the amount of clustering should indicate an increase in density. This did not occur enough times to make any inferences. In Bremer County, clustering increased in both study areas between 1875 and 1917, and density increased by about four percent in both areas. In Jones County, clustering increased between 1893 and 1915, but density decreased by about twelve percent. But the ratio changes are not significant. Those are the only cases where the hypothesis could be examined. Certainly no conclusions can be drawn.

Hypothesis 3 states that regularity should increase over time in a study area such as rural Iowa. There were twenty-seven cases in which the negative binomial fit a spatial distribution at some quadrat size with no significant differences for $\alpha = .05$ under the one-sample Kolmogorov-Smirnov test. Twenty of these occurred in the 19th century, five in the 1901–1917 period, and two occurred in 1960. The variance-mean ratio was significantly larger than one in six of the twenty-seven cases when it was larger than one. Five occurred in the first time period, one in the second, and none in 1960. The negative binomial yielded the best fit in twenty-one of the twenty-seven cases where it was possible to fit it. Fifteen were in the first, four in the second, and two were in the third time period.

The regular Poisson distribution was not significantly different from the data in sixty-eight cases, sixty-one excluding Cedar County since no comparable recent results were computed. Twelve occurred before 1900, twenty-five in the middle period, and twenty-four in 1960. The variance-mean ratio was significantly smaller than one in thirty-five cases; two in the first period, eleven in the second, and twenty-two in the third. The regular Poisson series was the best absolute fit all twelve times in the initial period, twenty-one times in the second, and twenty-four times in the third. In this last period four of the cases too extreme for this series occurred, also.[50]

These results indicate that the distribution is definitely tending toward greater and greater regularity. Over the roughly ninety-year period studied the number of farm residences declined by about five percent, although it was much greater in several study areas. These two facts, especially since density is no doubt overestimated for the most recent period, support hypotheses 1 and 3.

It is worth noting another interpretation, however. In eighty-three of the 105 cases, the Poisson series gave an adequate fit at the five percent level. In over three-fourths of the cases, the distribution's aggregative characteristics cannot be distinguished from a settlement process resting on chance, with no dependence between settlement locations on the map. However, this is not the hypothesis to be tested. In light of the theory of settlement process put forth here, it is more significant to note that although the Poisson series fit in eighty cases, it was as good or better than either alternative in only twenty cases: eight times in the first period, eleven in the second, and just once in the third. The early twentieth century was a period of transition between the higher-density more clustered settlement pattern of nineteenth century agriculture and the large-size more widely spaced farms of the present. Randomness is great in this middle period, as is expected. Thus, although the Poisson fits in many instances, the "richer," two-parameter distributions fit better, three-fourths of the time, giving support to the hypothesis that settlement process in this study area is not as simple as the Poisson distribution assumptions allow it to be.

One final point concerning goodness-of-fit involves the number of "no fits." In only two cases out of 108, do none of the three distributions adequately correspond to the data for $\alpha = .05$. Both frequency distributions were quite irregular. Such results indicate that these models alone are adequate to describe rural farm settlement distributions in an area such as Iowa. For this reason, the employment of any other models was felt to be redundant.

[50] The definition of the parameter, p, requires that $\bar{x} - 1 \leqslant \sigma_x^2 \leqslant \bar{x}$. This inequality is sometimes violated when the quadrat is large and the distribution highly regular. When this occurs, the regular Poisson series is undefined, but the limitation is not serious since the pattern is highly regular, anyway.

CONCLUSIONS

A theory of settlement location has been presented which attempts to explain the morphological changes that take place in a rural settlement distribution over time. When density is low, and unsettled areas are common, settlement locations are essentially independent of each other. As density increases through a continued diffusion of settlements, competition for space becomes increasingly important. The pattern changes from a clustered to a highly regular arrangement as weak individuals are forced out and the average size of holdings increase. The theory imposes conditions on the central place assumption of uniform base population, mainly by requiring sufficient competition for space to produce uniformity of spacing. Under this condition the size distribution of settlements must have a very small variance, if the settled area is physically homogeneous.

The validity of some of the theory was tested in actual settlement distributions in Iowa. It was found that regularity of spacing did increase with time, between 1870 and 1960, agreeing with theoretical expectations based on knowledge of the agricultural economy of the area during this time.

It should not be inferred that the theory of settlement location described here states that all distributions will pass through these pattern changes in precisely the same way and in the same length of time. Influences may shift from clustering to competition and back again, owing to changes in exogenous variables, here operationalized as parameters of the niche space. The role of government planning is an obvious omission from the model, which might cause just such a modification.[51] In such cases where this exists, the

[51] The critic may argue that the regular pattern of settlement found in Iowa is a function of the rectangular land survey system which creates a grid system of roads. However, even though the same

introduction of a planned settlement pattern that is rigidly followed makes the models presented here, superfluous. If the geographer knows the pattern to be one deliberately imposed, he has no need to invoke this theory to explain it. The existence of highly regular settlement distribution in carefully planned colonies is no more interesting or appropriate to this theory than the regular spacing of plants produced by a mechanical planter is to the theory of spatial competition in natural vegetation. The interesting cases are those where regular spacing occurs without such planning and instead occurs as the predictable result of several spatial processes whose interactions produce the distinctive geometry of regular forms. It is for these cases that the theory, hopefully, is appropriate.

roads and survey system existed throughout the period under examination, the settlement distribution changed from an irregular or clustered one, to a highly regular pattern. Also, regular settlement pattern may exist where such a survey system is not in existence. For example, see the map of Chinese agricultural settlements in J. W. Alexander, *Economic Geography* (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1963), p. 58.