TWO MODELS FOR THE INFERENTIAL ANALYSIS OF CENTRAL PLACE PATTERNS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

Clifford E. Tiedemann

1966

This is to certify that the

thesis entitled

TWO MODELS FOR THE INFERENTIAL ANALYSIS
OF CENTRAL PLACE PATTERNS

presented by

Clifford E. Tiedemann

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Geography

Donald A. Blome
Major professor

Date Aug 23, 1966

O-169

ABSTRACT

TWO MODELS FOR THE INFERENTIAL ANALYSIS
OF CENTRAL PLACE PATTERNS

by Clifford E. Tiedemann

Walter Christaller's central place theory has provided geographers with a logical construct which describes and explains the influences of relative location on the distributional patterns of cities. However, the theoretical statements by Christaller are predicated on a set of assumptions and rules which, it has been stated, are seldom found to be coincident in the observable world. Thus, assessments of real-world spatial patterns of agglomerated settlements in the light of this theory are extremely difficult.

Many studies have been made in which analyses of existing settlement patterns are based on selected aspects of central place theory. Although it must be admitted that the spacing of cities is an extremely complex object of study, these efforts are viewed as being partially unsatisfying, since they all fail to properly account for the distance -- population size relationship. It appears, then, a technique to be used in the evaluation of this association can be of some value in this field of endeavor.

Clifford E. Tiedemann

Using a stochastic process to define equality among populations of cities, two descriptive models are proposed which can account for this important factor in the spacing of settlements. The two models, one a simple regression and the second a simple correlation, indicate the nature of this bivariate relationship. The independent variable is settlement population, which serves as an indicator of the functional complexity of each central place. The dependent variable is standardized distance, a phenomenon which possesses characteristics of particular value in the analysis of the distribution of central places.

The parameters of these two models, the y-intercept, the regression coefficient, and the correlation coefficient, each have unique numerical values which are defined by the deterministic character of central place theory. Calculated values for each of these parameters can be compared with the defined figures using standard tests of significance based on the normal and "t" distributions. Relying on the outcome of such tests, statements of confidence can be made concerning the similarity, or lack thereof, between an observed settlement pattern and that theorized by Christaller. In addition, using only slightly modified tests of significance, it is possible to compare the respective parameters associated with two different settlement patterns.

Clifford E. Tiedemann

One thousand fifty four agglomerated settlements in Michigan with 1960 populations exceeding 100 are used in the demonstrations of the models. Analyses are reported in which the following comparisons are made: the regression coefficient of the entire state with that defined by central place theory; the regression coefficient of a subregion with that of the state; the regression coefficient of one subregion with that of another; the y-intercept of a subregion with that defined by central place theory; the y-intercept of one subregion with that of another; and the correlation coefficient of a subregion with that defined by central place theory. In each analysis, an underlying reason for the test is described and an hypothesis is offered, tested, and either accepted or rejected.

As a result of the development and demonstrations of the models, several terms are proposed with which the results of the tests of significance may be described. While many of these terms are refinements of previously used ideas, one is particularly interesting -- "differential clustering." This is a situation in which the regression coefficient is found to be significantly different from zero, indicating that one end of the array of populations of places shows a greater deviation from the uniform spatial distribution of central place theory than does the opposite end.

Clifford E. Tiedemann

Technical discussions are appended in which the stochastic process of determining equality among settlement populations and the calculation of metric and standardized distances are reviewed.

Approved: ______________________

Date: ______________________

TWO MODELS FOR THE INFERENTIAL ANALYSIS OF CENTRAL PLACE PATTERNS

by

Clifford E. Tiedemann

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Geography

1966

ACKNOWLEDGMENTS

The preparation of this dissertation has been aided by many individuals, and it is fitting that those who played key roles be cited. In order for proper credit to be given, names of persons are mentioned in groupings according to their activities associated with this work.

The first group consists of my academic advisory committee. Chaired by Dr. Donald A. Blome, this group of professors has acted in a most diligent manner in my behalf at all times. Also including Dr. Dieter H. Brunnschweiler, Dr. Harm J. deBlij and Dr. Milton H. Steinmueller of the Department of Resource Development, their quick replies and penetrating comments concerning my efforts are most appreciated. Also of assistance in this respect, but no longer at Michigan State University are Professors Allen K. Philbrick, now of Western Ontario University, and Julian Wolpert, now of the University of Pennsylvania.

I am particularly indebted to Drs. Blome and Philbrick for their influence on the conceptual framework of my research. These two gentlemen, with their respective interests in central place theory and areal functional organization, have done much to guide me in the selection of a field

of study and a problem. Without their willing and expert assistance, the undertaking of this project would have been at least considerably more difficult.

Technical assistance came from two colleagues in the Michigan Interuniversity Community of Mathematical Geographers. Dr. Waldo Tobler of the University of Michigan provided help in the computation of distances between the urban places used in the analyses. Anthony V. Williams, a fellow graduate student in geography at Michigan State University, provided much needed assistance where this writer's computer programming ability proved to be limited or faulty.

Finally, in a class by herself, I must recognize my wife, Margaret. Without her faith and constant support, my work as a graduate student would have been doomed from the beginning.

Each of the individuals mentioned deserves some credit for whatever good comes from this research. None of them can be held responsible for its shortcomings. My only hope is that someday I will be able to pass on to others the time and effort which they have invested in me.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iii

LIST OF TABLES vii

LIST OF ILLUSTRATIONS viii

Chapter

I. INTRODUCTION 1

Central Place Theory
Selected Studies Related to the Analysis of Central Place Patterns

II. DEVELOPMENT OF THE MODELS AND TESTS 21

The Conceptual Framework
The Basic Models
descriptive parameters
tests of significance
influences of error
Preliminary Evaluation

III. EMPIRICAL DEMONSTRATIONS OF THE MODELS 44

The Study Universe
Operational Definitions
Applications and Tests of the Models
the first test of significance
the second test of significance
the third test of significance

IV. CONCLUSION 81

Evaluation of the Empirical Applications of the Models
Evaluation of the Models and Tests

Page

APPENDIX A 91

APPENDIX B 99

BIBLIOGRAPHY 104

LIST OF TABLES

Table — Page

I-1. Selected Aspects of the Spatial and Functional Hierarchical Patterns of Central Place Theory 5

III-1. Number of Settlements with Populations Equal to or Larger than Selected Settlements 51

III-2. Expected Distances Between Selected Settlements 52

LIST OF ILLUSTRATIONS

Figure		Page
II-1.	Expected Distributional Relationships	30
II-2.	Expected Distributional Relationships	33
III-1.	Fractile Diagram: Test for Normality (Populations)	50
III-2.	Fractile Diagram: Test for Normality (Standardized Distances)	54
III-3.	Michigan: Settlements with More Than 100 Persons	56
III-4.	The Size-Distance Relationships for Settlements with More Than 100 Inhabitants . . . A. For All of Michigan B. For the Western Part of the Upper Peninsula C. For the Lower Peninsula i. inland counties ii. shoreline counties	58
III-5.	Michigan: Settlements with More Than 100 Persons in the Western Part of the Upper Peninsula	61
III-6.	Michigan: Settlements with More Than 100 Persons in the Shoreline Counties of the Lower Peninsula	64
III-7.	Michigan: Settlements with More Than 100 Persons in the Inland Counties of the Lower Peninsula	65
III-8.	Michigan: Settlements with More Than 100 Persons in State Economic Areas 6 and 7 . .	70

LIST OF ILLUSTRATIONS (Continued)

Figure | Page

III-9. The Size-Distance Relationship for Settlements with More Than 100 Inhabitants . . . 71
- D. For State Economic Area No. 6
- E. For State Economic Area No. 8 and No. 9
  - i. S.E.A.-8
  - ii. S.E.A.-9
- F. For State Economic Area No. 7

III-10. Michigan: Settlements with More Than 100 Persons in State Economic Areas 8 and 9 . 74

CHAPTER I

INTRODUCTION

The distribution of phenomena over the earth is of primary interest to the geographer. One such phenomenon is the city,[1] and Walter Christaller's theoretical formulation of the distribution of agglomerated settlements in Southern Germany has done much to direct the interests and efforts of several contemporary geographers and students of other disciplines.[2] Since the appearance of this work, many studies have been conducted and papers written on topics closely related to central place theory.[3]

Central Place Theory

In developing the conceptual framework with which Christaller constructs central place theory, he notes that

---

[1]The words "city," "town," "settlement," and other terms with connotations of human population agglomeration are used interchangeably for purposes of ease of expression. If a particular definition is required at some point in the text, it is provided at the appropriate place.

[2]Walter Christaller, Central Places in Southern Germany, Carlisle W. Baskin, trans. (Englewood Cliffs: Prentice Hall, Inc., 1966).

[3]Brian J. L. Berry and Alan Pred, Central Place Studies: A Bibliography of Theory and Applications (Philadelphia: Regional Science Research Institute, 1961).

there exist certain goods and services which are to be acquired only at specific locations. But, the demand for these items is found throughout the population, regardless of the spatial distribution of individuals. These goods and services he calls "central goods and services," and the locations at which they are available are designated "central places."[4]

Each central good or service, in order to be made available in any region, must have a demand of sufficient amount to support the necessary marketing activities. A central place, then, is associated with a "complementary region" of a size which contains that number of people capable of generating an aggregate demand such that a profit is made in the sale of the good or service.[5] The extent of a complementary region is limited, however, by the cost of the transportation involved in procuring a good or service, and by the relative ease of accessability to another central place offering a similar good or service. Before a particular central good or service will be made available to consumers at a given place, it is axiomatic that the complementary region be of sufficient area to enclose the minimum required demand for the relevant item.

---

[4]Christaller, loc. cit., pp. 14-21.

[5]Ibid., pp. 21-2 and 60.

The function of a central place is "to be [the] center of its rural surroundings and mediator of local commerce with the outside world."[6] The locations of central places, then, are determined to be relative to the region which they serve. Such a locational definition implicitly eliminates from consideration as central places all those settlements whose sites are determined to be specifically oriented to some phenomenon. Christaller explicitly lists such nucleated settlements as mining towns, border and ford sites, and towns at other unique locations as monasteries and shrines, and even residential suburbs of large industrial urban centers.[7] It is also pointed out that among those towns which may be considered as central places, "there is a definite connection between the consumption of central goods and the development of those central places. The development of those central places whose inhabitants live by the sale of central goods becomes more pronounced if many central goods are consumed than if few central goods are consumed."[8]

Using these terms and relationships, Christaller develops a theory with which he explains the spatial and functional hierarchical patterns of central place distribution.

---

[6] Ibid., p. 16.

[7] Ibid., pp. 16-7.

[8] Ibid., p. 27.

He does so using an idealized environmental situation based on the following rules and assumptions:

(1) a flat, unbounded plain;

(2) a homogeneous resource base;

(3) a ubiquitous transportation system;

(4) an even distribution of population with uniform, individual purchasing power;

(5) the sale prices of central goods of any one type are constant through time and uniform over space;

(6) all demands for central goods are satisfied, provided the seller makes a profit;

(7) profits for any one seller cannot equal or exceed that amount which can support two such activities, and

(8) any central place offering good i must also offer all goods j which require equal or smaller minimum demands.

Within this framework, rigid spatial and functional hierarchical patterns of growth and development would occur (see Table I-1). Indeed, Christaller concludes his initial formulations as follows:

> The result of these theoretical considerations is surprising but clear. First, the central places are distributed over the region according to certain laws. Surrounding a greater place (B-type), there is a wreath of the smallest places (M-type). Furthermore, there is another wreath of the small places (A-type). Towards the periphery there are a second and third wreath

Table I-1. Selected Aspects of the Spatial and Functional Hierarchical Patterns of Central Place Theory[9]

Order Class (Type)	Number of Places	Numbers of Complementary Regions	Range of Central Good (km.)	Area of Region (sq. km.)	Typical Populations of Central Places	Typical Populations of Regions
M	486	729	4.	44	1,000	3,500
A	162	243	6.9	133	2,000	11,000
K	54	81	12.	400	4,000	35,000
B	18	27	20.7	1,200	10,000	100,000
G	6	9	36.	3,600	30,000	350,000
P	2	3	62.1	10,800	100,000	1,000,000
L	1	1	108.	32,400	500,000	3,500,000

[9] *Ibid.*, p. 67.

> of the smallest places (M-type), and on the periphery itself we find the middle-sized places of the K-type. The same rules are valid for the development of greater systems. Second, following the laws of economics, there are necessarily quite definite size-types of central places as well as the complementary regions, and indeed characteristic types, not order classes. Third, the number of central places and their complementary regions which are to be counted for every type form a geometric progression from the highest to the lowest type.[10]

Thus, in his original statement, Christaller has formulated a conceptual model of the patterns of town development which is clearly deterministic in character. That is, under the conditions specified by the rules, assumptions and relationships, certain regular spatial and population-size patterns develop.

As can be seen in Table I-1, the populations of central places of different types in the functional hierarchy have "typical" magnitudes. The discreteness of the distribution of these typical populations is associated with the relationship between the functional complexity of a central place and its population. And, since every central place offers the entire range of central goods and services requiring a demand equal to or less than its highest order item, a step-like hierarchy of central places with a step-like distribution of populations develops under the specified conditions.

---

[10] _Ibid._, pp. 66-8.

These conditions also lead to the formation of a recognizeable spatial pattern--the regular hexagonal lattice. In the instance of Table I-1, the lattice contains 486 points, each representing the location of a central place of one type or another. All of the places are centers for the lowest orders (M-type) of goods, and in the example are located approximately eight kilometers apart, or twice the range of that order good. Of the 486 M-type central places, 162 also handle the next higher order of central goods (A-type). These A-type central places are also distributed hexagonally with respect to one another, and since they share locations with some of the M-type places, they are located 13.8 (or $2 \cdot 4 \cdot \sqrt{3}$) kilometers apart. The next higher type of central places (K-type) share locations with fifty-four of the A- and M-types of places. The distance between neighboring K-type places is twenty-four (or $2 \cdot 6.9 \cdot \sqrt{3}$) kilometers, since they, like the A and M places, are arranged in a regular hexagonal pattern. Similar relationships are apparent in the spacing of all central places as one moves through the entire functional hierarchy. Within the context of central place theory, then, it can be noted that if the distance between two neighboring towns of the same functional order is d, then the distances between neighboring places both of which are m levels above or below the original pair in the hierarchy are $d(\sqrt{3})^m$ or

$d(\sqrt{3})^{1/m}$ respectively.[11] Thus, a second discrete distribution is produced by Christaller's idealized specifications.

Table I-1 illustrates that as the "typical" population increases in its step-wise fashion, so does the distance between neighboring central places of the same functional type. It is this distance-population-size relationship which forms the basis of this research.

Selected Studies Related to the Analysis of Central Place Patterns

One of the earliest discussions of the distribution of cities over space which is based to a great degree on central place theory is that by Lösch.[12] Lösch assumes areal conditions quite similar to those of Christaller, and for any system of places offering a particular central good, he generates an hexagonal lattice of central place locations and complementary regions. Although it is not readily apparent in this work, it becomes obvious in a subsequent book that Lösch does not require that any central place offering good i must also offer all goods j which require smaller minimum demands.[13] Thus, Lösch's manipulations of

---

[11] Ibid., p. 63.

[12] August Lösch, "The Nature of Economic Regions," Southern Economic Journal V, 2 (July, 1938), pp. 71-8.

[13] August Lösch, The Economics of Location, W. H. Woglom and W. F. Stolper, trans. (New Haven: Yale University Press, 1954), pp. 118-9.

the hexagonal trade areas (complementary regions) produces a regular hexagonal lattice with a spacing pattern for cities of similar functional complexity which is different from that of Christaller. Isard criticizes this construct as being inconsistent with Lösch's own assumptions, since it produces regions with different concentrations of cities and, therefore, of people.[14]

An early application of central place theory in America was performed by Brush in his study of settlements in Southwestern Wisconsin.[15] In this effort, Brush recognizes three types of settlements -- hamlets, villages and towns -- each of which is characterized by its degree of functional complexity. He then compares the spacing of the various centers with that which would be expected if a regular hexagonal pattern were present, and his average distances separating places with equal status in the functional hierarchy are quite close to those which would be generated by central place theory.

Criticism of Brush's work is addressed to both his settlement classification scheme and his evaluation of the

---

[14]Walter Isard, Location and Space Economy (New York: John Wiley and Sons, Inc., 1956), p. 271.

[15]John E. Brush, "The Hierarchy of Central Places in Southwestern Wisconsin," Geographical Review XLIII, 3 (July, 1953), pp. 380-402.

spatial distribution of the central places. Vining, pointing out that Brush has not been able to demonstrate a step-like hierarchy of settlement populations, bases his remarks primarily upon the three-fold grouping of settlements.[16]

> . . .After having specified the criteria by which an enumerator may distinguish among various kinds of activities, the observer has as his basic data the array of communities and the listing of the kinds of activities represented by the establishments in each community. There is no evidence that I have seen suggesting that exactly three natural partitions may be observed in this array of numbers of establishments. Like pool, pond and lake, the terms hamlet, village and town are convenient modes of expression, but they do not refer to structurally distinct natural entities. As the number of establishments increases, the number of kinds of activities represented also increased. Clearly it is arbitrary to divide the array into three partitions rather than into a greater or lesser number; and similarly arbitrary is the determination of where to put the dividing points separating the different classes or types. Having drawn the lines, one may list certain kinds of activities which are typically found within each of the designated classes of center, and . . . [the table in] . . . Mr. Brush's article represents such a listing. It will be noted that not all members of a class will contain all the activities listed, and most of the communities within a class will contain activities not listed. Such a table is not an independently derived basis for a classification of communities by type. Rather it is itself derived from a previous partitioning of an array which appears as something similar to an arrangement of observations that have been made upon a continuous variable.

---

[16]Rutledge Vining, "A Description of Certain Spatial Aspects of an Economic System" Economic Development and Cultural Change III, 2 (January, 1955), p. 160.

Such a criticism, of course, indicates a need for a more rational scheme of city classification in the instance where a relationship between functional complexity and population is to be considered.

In a discussion of Vining's paper, Hoover mentions a logical scheme for evaluating the functional and spatial relationships of an observed group of cities within the framework of central place theory.[17] This approach takes into account the rank-size of a city and its link with the functional complexity of the city.

> What is it exactly that makes the series of tributary areas of cities follow the rank size rule? It is convenient at this point to think of the series as continuous rather than discrete. In other words, each city is in a class by itself, and performs some function that is performed by <u>no</u> smaller city, but by <u>all</u> larger cities. Now if we take, say, the 17th biggest area, we can consider it as <u>one of the group of the 17 biggest</u>. With respect to the function exclusively shared by these 17, the whole country is parceled out into 17 equal areas. Similarly the 16th biggest is one of 16 equal areas blanketing the country with respect to the function which <u>only the 16 biggest</u> cities performed. Obviously, the rank-size progression is implicit, then, in the concept of each city's area being determined by equal sharing with all larger cities.[18]

---

[17] Edgar M. Hoover, "The Concept of a System of Cities: A Comment on Rutledge Vining's Paper," <u>Economic Development and Cultural Change</u> III, 2 (January, 1955), p. 197.

[18] For a discussion of several rank-size theories, see Brian J. L. Berry and William L. Garrison, "Alternate Explanations of Urban Rank-Size Relationships," <u>Annals</u>, Association of American Geographers XLVIII, 1 (March, 1958), pp. 83-91.

It is apparent in this statement that Hoover is addressing himself at least in part to the spatial aspects of the problem. But, it is to be remembered that according to Christaller's concepts, a direct relationship exists between

(1) the size of the area being served by a town and the population of the area,

(2) the population of the area being served and the functional complexity of the central place, and

(3) the functional complexity of a central place and its population.

Thus, it is conceivable that a continuous distribution of population is amenable to an analysis based on central place theory, without resorting to the arbitrary groupings criticized by Vining. But this scheme is very restrictive, in that it allows no consideration to be given to places even slightly smaller than a given settlement.

In a series of short articles, Berry and Garrison attempt to relate the conditions and constructs of central place theory with observed situations and patterns in the real world.[19] Relying on the concepts of the range of a

---

[19] Brian J. L. Berry and William L. Garrison, "The Functional Bases of the Central Place Hierarchy," Economic Geography XXXLV, 2 (April, 1958), pp. 145-54.

Brian J. L. Berry and William L. Garrison, "A Note on Central Place Theory and the Range of a Good," Economic Geography XXXIV, 4 (October, 1958), pp. 304-11.

good and the threshold -- both of which are consistent with Christaller -- these researchers first recognize problems and possible applications of the principles of central place theory. Ultimately many of the more restrictive conditions are relaxed, and, it is reported, a nesting of central places is found to exist. It is also learned that certain central goods require different threshold populations, and a functional hierarchy of cities is formulated based on the numbers and kinds of activities present in cities of various populations. Thus, it is demonstrated that a functional hierarchy of central places will develop under conditions less restrictive than those outlined by Christaller, and that this functional hierarchy is closely related to the distribution of settlement populations.

Thomas assumes the presence of an association between city size and spacing, and demonstrates that this relationship has not changed significantly during the period 1900-1950 in Iowa.[20] Using inferential statistical methods on a sample of eighty-nine Iowa cities, the author finds that

---

Brian J. L. Berry and William L. Garrison, "Recent Developments of Central Place Theory," Papers and Proceedings, Regional Science Association IV (1958), pp. 107-20.

[20]Edwin N. Thomas, "The Stability of Distance-Population-Size Relationships for Iowa Towns from 1900-1950," Proceedings of the I. G. U. Symposium in Urban Geography, Lund, 1960: Lund Studies in Geography, Series B. Human Geography, No. 24 Knut Norborg, ed. (Lund: The Royal University of Lund, Sweden, Department of Geography, 1962), pp. 13-30.

a positive correlation exists between the population of a given city and the distance to its nearest neighbor of equal or larger size for the period under study. This, of course, is consistent with central place theory in that increasing population is associated with increasing distance.

In this research, Thomas uses a stochastic model to expand upon the definition of "equal size," setting up requirements which he discusses in a subsequent report.[21] He states:

> . . .we do not mean that the population of the sample city and the neighbor city are the same; that is, we do not mean that
>
> $$S_i = N_1 \qquad (1)$$
>
> where S_i is the population of the i^{th} sample city and N_i is population of its nearest neighbor. However, we do mean that S_i is approximately equal to N_i; that is
>
> $$S_i \approx N_i \qquad (2)$$
>
> In addition we will say that
>
> $$S_i + R_i = N_i \qquad (3)$$
>
> where R_i is a variable whose magnitude is due to chance; that is, R_i is a stochastic "error" variable. The phrase "same population size" is defined in terms of (3), and, accordingly, we may say that populations are accepted as having the same size when they differ only by chance.[22]

---

[21] Edwin N. Thomas, "Toward an Expanded Central Place Model," Geographical Review LI, 3 (September, 1961), pp. 400-11.

[22] *Ibid.*, p. 403.

Thus, a relationship is established between the discrete and typical size features of the distribution of populations according to central place theory and the problem of the continuous distribution of settlement sizes as recognized by Vining and Hoover.

Using similar definitions of equal population size for towns, King and Blome expand upon the basic analytic model as described by Thomas.[23] Both of these works use multivariate statistical techniques in attempts to explain the spatial patterns of settlement distribution. King's effort is aimed at detecting the most important of those variables which he contends influence the spacing of cities at a particular time -- the census year, 1950. Blome, on the other hand, traces the changing relationships between a variety of factors and city spacing over a period of sixty years. Both of these studies, as do those of Thomas, acknowledge that the spacing of settlements is closely related to the dynamic associations between towns and their surroundings, including both contiguous rural and urban developments and similar urban phenomena located some distance away.

---

[23]Leslie J. King, "A Multivariate Analysis of the Spacing of Urban Settlements in the United States," Annals, Association of American Geographers LI, 2 (June, 1961), pp. 222-33.

Donald A. Blome, "An Analysis of the Changing Spatial Relationships of Iowa Towns, 1900-1960," unpublished Ph.D. dissertation (Iowa City: The State University of Iowa, Department of Geography, 1963).

Dacey approaches the spatial aspects of city distribution from a point of view quite different from that of many geographers. His work with central place theory is based primarily upon an interest in the distributional aspects of point patterns through spaces of various dimensions. Much of Dacey's work is based upon nonparametric statistics, with which he describes the spacing of points in terms of their number and the area in which they are distributed.

In a discussion of Brush's work in Southwestern Wisconsin, Dacey assumes that the defined classification scheme is valid.[24] He then proceeds to demonstrate that the centers of each level of the hierarchy are distributed according to a random pattern. This, of course, does not conform to the regular hexagonal construct of central place theory, and it does not agree with the apparent close comparison illustrated by Brush.[25]

In his analysis, Dacey reduces the observed distances between pairs of settlements in Southwestern Wisconsin to a common scale, taking into account the density of the

---

[24] Michael F. Dacey, "Analysis of Central Place and Point Patterns by a Nearest Neighbor Method," Proceedings of the I. G. U. Symposium in Urban Geography, Lund, 1960: Lund Studies in Geography, Series B. Human Geography, No. 24 Knut Norborg, ed. (Lund: The Royal University of Lund, Sweden, Department of Geography, 1962), pp. 55-76.

[25] Brush, loc. cit., p. 393.

places being studied. He refers to this reduced value as the "standardized distance," and to the reduction factor as a "dimensional constant."[26]

> The measured map distance from i [a point] to the nearest point (j) is represented by R_{ij}... The R_{ij} measurements reflect the arbitrary map metric. The dimensional constant which eliminates effect of scale is $d^{1}/2$, where d is the density of points in Q [an area]. Measurements in Q are reduced to standardized distance by the transformation
>
> $$r_{ij} = d^{1}/2R_{ij}$$

Thus, the distances between pairs of points -- in this instance, cities -- are rendered dimensionless, and are significant only as to their relative magnitudes. That is, r_{ij} has no metric meaning, and will be the same for any one observed distance regardless of whether R_{ij} and Q are measured in terms of miles and square miles, kilometers and square kilometers, or even inches and square inches respectively.[27]

This measure of standardized distance, however, has an additional property which is most useful, and is what formed

---

[26]Michael F. Dacey, Imperfections in the Uniform Plane: Discussion Paper No. 4 (Ann Arbor: University of Michigan. Department of Geography for the Michigan Interuniversity Community of Mathematical Geographers, June, 1964). (mimeo.)

[27]The density of points, d, in an area, Q, is obtained by the following equation:

$$d = n / Q,$$

where n is the number of points in Q.

the basis for Dacey's analysis of Brush's work. If in any given area, the places under consideration are arranged in a regular hexagonal manner, all of the r_{ij}'s will have a common value, say c. Marked deviations from a uniform spatial distribution, however, will produce a mean standardized distance, $\bar{r}_{ij}$, which will be smaller than c. Tests may be performed in which $\bar{r}_{ij}$ is compared with an <u>expected</u> mean standardized distance, e_{ij}, which is associated with a random distribution of points having the same density. The results of these tests allow the researcher to arrive at results to the effect that a given point distribution is "more uniform than random," "not significantly different from random," and "more clustered than random."

All of the models described above fail to account for certain relationships which are of considerable significance in studying the distribution of settlements within the framework of central place theory. A few statements will suffice in relating these problems.

(1) Brush's work was criticized because of the manner in which he devised his functional classification scheme which he then used to define his hierarchy of central places.

(2) Hoover accepts the reality of continuous distribution of settlement populations, but is extremely restrictive in his method of handling

the problem of functional (population-size) equality.

(3) Thomas expands upon Hoover's notions, but in his attempts to relate distance to population, he constructs a model which fails to take into account the spatial relations between cities and the area in which they are distributed.

(4) King and Blome, despite their impressive explanatory models, also do not account for the nature of point distributions, thus rendering their results not comparable with similar studies which might be made elsewhere.

(5) Dacey recognizes the spatial dynamics of point distributions, but does not consider anything other than a situation in which all points have a unit value. He also fails to provide a method for the analysis of differentials in distribution characteristics either over space or through a hierarchy of point classes.

The object of this research is to develop a model or group of models which are capable of providing for these needs.

What is proposed and demonstrated in the following chapters is a group of models and tests which are capable of doing the following:

(1) describe the relationship between population-sizes of cities and their spacing;

(2) test this relationship for:

  (a) conformity to central place theory over the entire study area,

  (b) conformity to central place theory over less than the entire study area,

  (c) a random distribution of cities of all sizes over the entire study area,

  (d) differential clustering among cities of various population sizes;

(3) test for significant differences between the size-distance parameters for different groups of cities;

(4) provide a basis for explanatory models which can account for differences in spatial patterns of settlement distribution.

Such a group of models can be of great value in analyzing and describing a wide variety of situations and influences relevant to the spacing of agglomerated settlements over the earth's surface. They derive their usefulness from the fact that they can be used both to test deviations from theoretically-based "expected conditions," and to compare observed patterns.

CHAPTER II

DEVELOPMENT OF THE MODELS AND TESTS

Walter Christaller, with his conceptual statements concerning the role of economic centrality in the development of settlement patterns, has provided students of such phenomena with a rigorous theory relating cities to their rural surroundings and to one another. Central place theory takes into account both population-size and distance between similar neighboring towns, and furnishes a logically constructed, "expected" situation with respect to these two variables. Indeed, within the context of the theory, this relationship between size and spacing is deterministic in character, there being matched step functions for both variables and interdependent, paired observations for every settlement.

Dacey characterizes central place theory as being ". . . a deductive formulation from very restrictive conditions . . . . a flat, unbounded, homogeneous area, an evenly distributed population density, and a ubiquitous transportation system."[28] He then observes that these conditions are seldom found to

---

[28] Michael F. Dacey, "Analysis of Central Place and Point Patterns. . .," loc.cit., p. 59.

be coincident in reality. However, it is interesting to note in his criticism of Brush's work that Dacey recognizes five point pattern relationships involving settlement function and spacing and compares the observed respective spatial distributions with those which would be expected to develop if central place theory's idealized conditions prevailed. Thus, it appears that the size-distance association is one phenomenon which can be used to analyze the distribution of agglomerated settlements -- even if Christaller's specifications are not to be found.

The Conceptual Framework

In order to achieve consistency with central place theory, it is necessary to take into account certain qualities of the two variables involved in the analysis of the size-distance relationship.

Vining's criticism of Brush, as discussed earlier, points out one of two basic problems which one encounters in attempting to analyze a settlement pattern within the context of Christaller's theory. Namely, this problem is centered about the categorization of cities by size and functional complexity, especially in the instance where the distribution of settlement populations is continuous and that of functional complexity appears to be somewhat jumbled, though grossly related to population.[29] Hoover's comment on Vining's criticism offers a partial solution, but it is too restrictive to be of great value.[30] That is, his qualification, "no smaller city," does not conform to central place theory, since even Christaller noticed groupings of populations about certain discrete values rather than absolute equality among city sizes within any one functional category.

Using a technique known as the "fractile diagram," Thomas expands upon Hoover's contribution, and establishes a link between the discrete and typical size features of central place theory and the problem of the continuous distribution of settlement populations as recognized by Vining and Hoover.[31] According to this method, there exist as many "equal size" categories as there are different town populations. And, instead of dividing a region into progressively smaller areal units as we move iteratively through the rank scale, as Hoover suggests, we now divide it into as many units as there are settlements with populations

---

[29]Rùtledge Vining, loc. cit.

[30]Edgar M. Hoover, loc. cit.

[31]Edwin N. Thomas, "Toward an Expanded. . .," loc. cit.

approximately equal to or larger than any given town, regardless of its rank.[32]

The fact that Thomas' proposed technique renders a continuous distribution of settlement populations amenable to central place theory, however, forces a second problem into view. That is, since the step-function of city sizes has been generalized, the size-distance relationship is no longer clearly defined in the theory's deterministic terms.

A solution to this question is contained in Dacey's criticism of Brush.[33] In the work, he points out that the distance between center points of two regular hexagons having one edge and two angles in common is

$$r = 1.075 \cdot \sqrt{H},$$

where r is the distance between the center points and H is

---

[32]The use of fractile diagram for the purpose of establishing equal size categories requires that the population data be of a normal form. (In addition to providing these categories, the fractile diagram is also an excellent test for normality, see Appendix A.) This requirement, of course, may necessitate some type of transformation of the data in order to render it compatible with the assumptions required by a particular statistical technique. The equal-size categories which are calculated then, are in terms of the transformed data rather than in the mode of the observations. Also, it should be noted that the resultant categories may not be mutually exclusive for places with similar populations.

[33]Michael F. Dacey, loc. cit., p. 62.

the area of each of two hexagons. In a large region which is completely covered with regular hexagons, the area of each is

$$H = A / n$$

where A is the area of the region and n is the number of regular hexagons it contains. The dimensional constant which removes the map metric from consideration, however, is the square root of the density, d,

$$\sqrt{d} = (n / A)^{1/2}.$$

Multiplying this conversion factor by R, an observed distance, produces a standardized distance

$$\begin{aligned} r &= (1.075 \cdot \sqrt{H}) \cdot \sqrt{d} \\ &= 1.075 \cdot (A / n)^{1/2} \cdot (n / A)^{1/2} \\ &= 1.075. \end{aligned}$$

This relationship, r = 1.075, holds constant regardless of the number of centers and hexagons present in a region.

The distance step function multiplier, $\sqrt{3}$, is not operative, then, when evaluating the spacing of central places with the use of standardized distances. Instead, the standardized distances associated with any level in an ideal central place hierarchy are equal to the constant value, 1.075. Also, 1.075 is the value of r for any hexagonal lattice, regardless of the number of regular hexagons. Standardized distance, therefore, may be used to evaluate a continuous distribution of the spacing of comparable central places even in areas where a uniform distribution of locations is not to be found.[34] This is so, because for any pair of places, if they and all comparable settlements were arranged according to a regular hexagonal lattice, the associated value of r will be 1.075.

The Basic Models

Descriptive Parameters

Assume that the distance separating a given settlement from its nearest neighbor having an equal or larger population (as discussed earlier) is dependent upon the population of the given place. This relationship is consistent with central place theory, and is evident in certain of the studies reviewed in the previous chapter. The nature of this relationship may be expressed in the form of a simple linear regression equation

$$y_i = a + bx_i, \tag{1}$$

---

[34] Uniformity in spatial distributions requires that two conditions be satisfied.

(1) individual points are evenly spaced from one another, and

(2) a maximum number of points must be packed into any small unit area.

These specifications in two dimensional space define a regular hexagonal lattice of point locations.

where y_i is the standardized distance to the nearest neighbor of equal or larger size associated with place i, x_i is the population of place i, and a and b are parameters which describe the linear relationship between the two variables. Indeed, if both the dependent and independent variables are normally distributed, this model can be easily tested for statistically significant deviations from various expected relationships in several ways.

A third descriptive parameter, the correlation coefficient, is a measure of how well individual observations conform to the regression model. This index, r, has the value

$$r = b \cdot s_x / s_y, \tag{2}$$

where s_x is the standard deviation of the independent variable (population) and s_y is the standard deviation of the dependent variable (standardized distance). The value of r can also be tested for significant differences from specified values.

Tests of Significance

The tests of significance which may be utilized in the analysis of these descriptive models are drawn from the realm of inferential statistics. In order to be applied, they require that certain conditions be met. Of primary importance, is the necessity that the distributions of the independent and dependent variables be bivariate normal.

In order to conform to the specified distributions, it may be necessary to transform one or both of the observed distributions into a form which more nearly approximates normality.

Two other conditions are concerned with the nature of the relationship between the two variables, rather than the characteristics of the distributions themselves. These requirements state that

(1) the standard error of the estimate must be constant for all values of the independent variable, and

(2) the distribution of the relevant dependent variables is normal for any given value of the independent variable.

If all of the conditions stated above are met, then tests of significance may be executed involving the descriptive parameters, a, b and r.[35] Based on these tests, statements of confidence may then be made concerning the characteristics of the size-distance relationship for the cities under investigation.

The descriptive parameters are products of the descriptive models (1) and (2). The tests of significance

---

[35]Wilfrid J. Dixon and Frank J. Massey, Jr. *Introduction to Statistical Analysis* (New York: McGraw-Hill Book Co., Inc., 1957), pp. 193-201.

involving the descriptive parameters, on the other hand, are analytic devices. The difference, of course, lies in the uses of these two types of formulations. One type, the descriptive model, illustrates one or more characteristics of the relationship between the independent and dependent variables. The second type, the analytic tool, involves the comparing of equivalent descriptive parameters from two relationships, or the comparing of one such parameter from a given relationship with that of an expected situation.

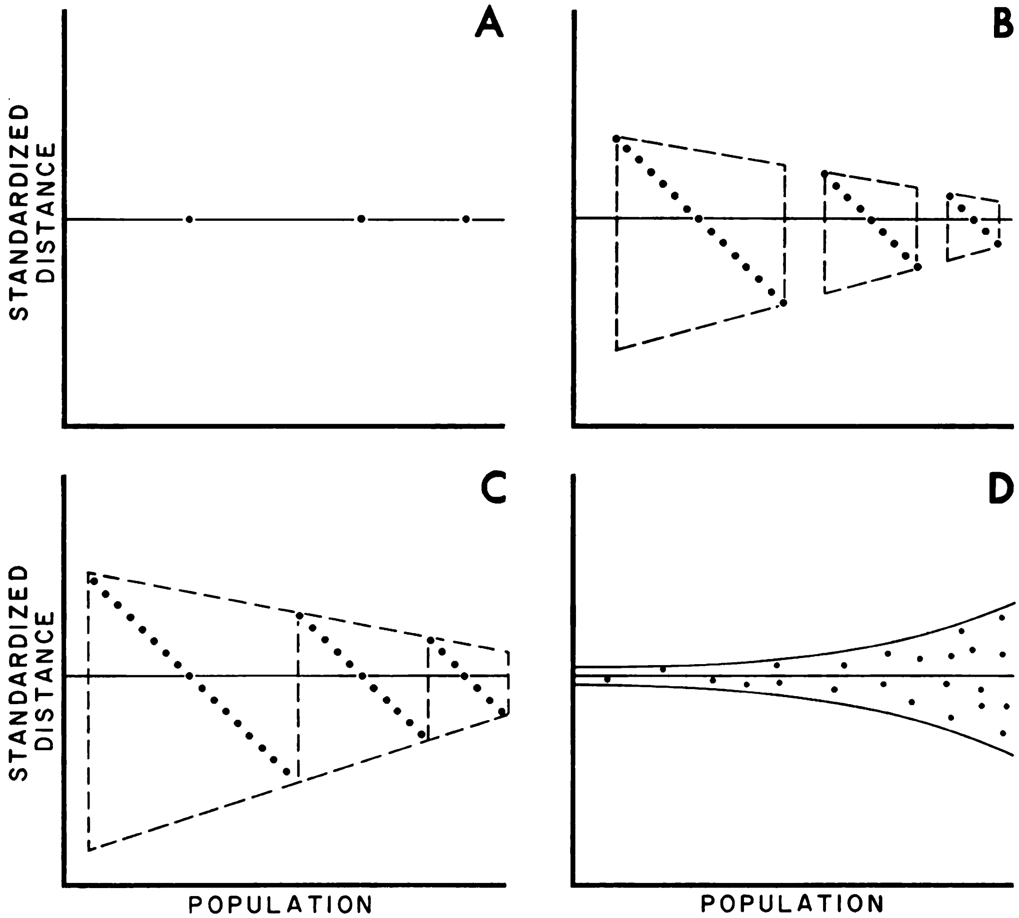
In the instance of central place theory, the expected situation is predefined. Figure II-1 shows the relationship between population size and standardized distance for a group of cities distributed according to a regular hexagonal lattice covering an entire study area. In the first case (see Figure II-1A), the idealized central place pattern -- that is, one having step functions for both city size and spacing -- is illustrated by a horizontal regression line. The regression equation for this is

$$\begin{aligned} y_i &= a + bx_i \\ &= 1.075 + 0.0x_i \\ &= 1.075. \end{aligned}$$

Since there is no deviation from 1.075 for any value of y, the correlation coefficient is

$$r = 1.0.$$

Fig. II-1. EXPECTED DISTRIBUTIONAL RELATIONSHIPS[36]



(A) an idealized central place pattern with the theorized discreet distribution of populations
(B) with populations clustered about "typical" magnitudes
(C) with a continuous distribution of populations
(D) a random dispersal over the prescribed locations

---

[36]assuming a regular hexagonal lattice of central place locations

In the second and third examples (see Figure II-1B and C), the restrictive step function of populations is replaced by clusters about "typical" size values and by a continuous distribution of populations, although the predefined locational pattern is retained. In both relationships the regression equation has the same parameters as those of the idealized central place patterns. The correlation coefficient, however, is

$$\begin{aligned} r &= b \cdot s_x / s_y \\ &= 0.0 \cdot s_x / s_y \\ &= 0.0, \end{aligned}$$

since the values of y vary about 1.075. Thus, if the slope of the regression line, b, is exactly zero, the correlation coefficient cannot be used as an indicator of conformity of an observed settlement pattern to that specified by central place theory.

In the fourth example (see Figure II-1D), the hexagonal pattern of place locations is retained, but the distribution of cities of various sizes among the locations is random. Again, the regression equation is the same as that for the previous three relationships, having a b-value of zero and a y-intercept, a, of 1.075. As in the cases of Figures II-1B and C, the correlation coefficient, r, is zero. The fourth example, of course, is not consistent with central place theory, but resembles the theoretical patterns of settlement location and possesses similar values for the descriptive parameters.

The analytic models are designed to test for significant differences between observed and expected size-distances relationships. That is, the observed value for b is compared with the expected value (zero), the observed a-value is compared with 1.075, and the observed correlation coefficient is compared with unity. Significant differences between the observed values of these descriptive parameters and those dictated by the theory indicate the presence of deviations from central place theory in the distribution of settlements within a study area or group of cities (see Figure II-2).

The first test involves the slope of regression line, b. This analytic model is used herein to detect the presence of a bias in the size-distance relationship which is not consistent with central place theory. The test statistic, t_b, has the value

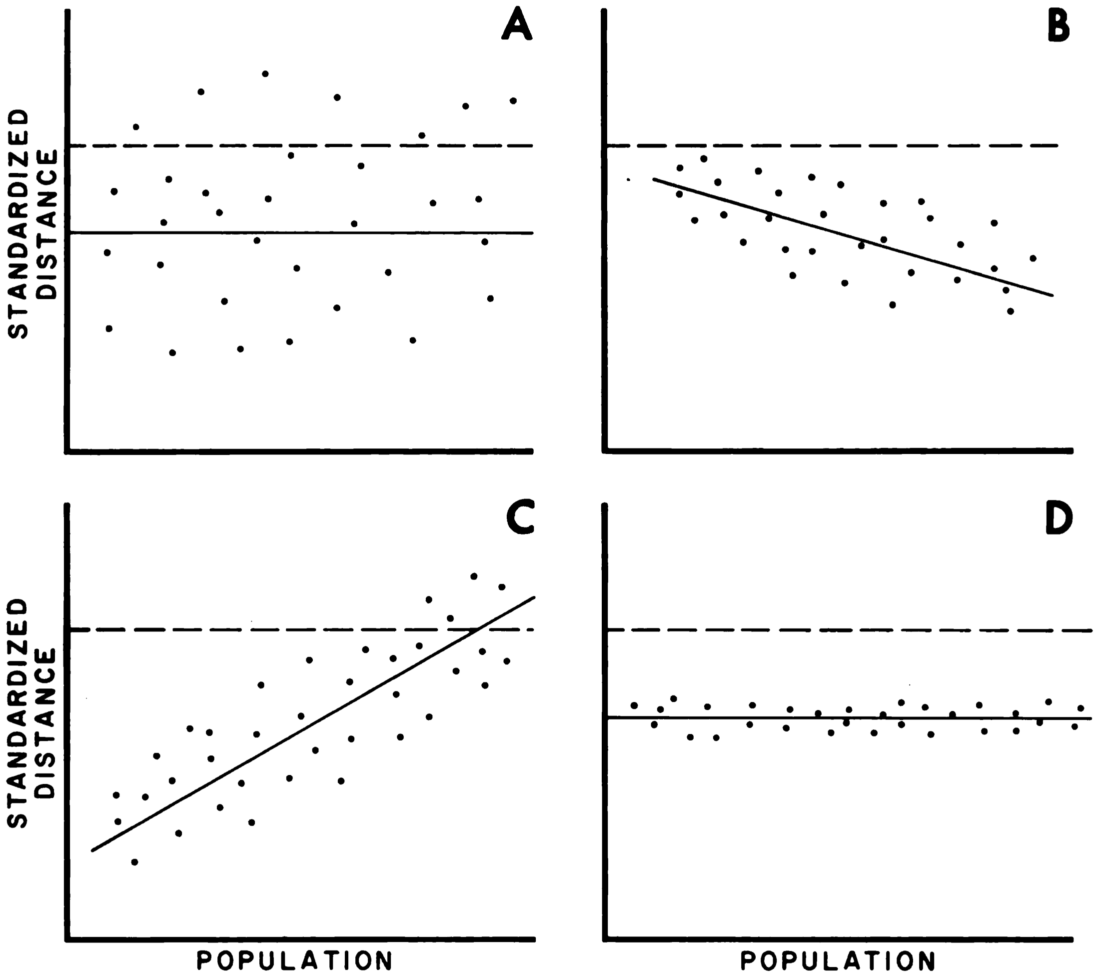
$$t_b = ((b - 0) \cdot s_x \cdot \sqrt{N - 1}) \, / \, s_{yx}, \qquad (3)$$

in which b is the slope of the regression line in the descriptive model

$$y_i = a + bx_i,$$

where 0 (zero) is the expected slope of the regression line according to central place theory, s_x is the standard deviation of the independent variable, N is the number of

Fig. II-2. EXPECTED DISTRIBUTIONAL RELATIONSHIPS[37]



(A) a random spatial pattern
(B) increased clustering with increased population
(C) decreased clustering with increased population
(D) an idealized central place pattern covering less than the entire study area

---

[37]assuming a continuous distribution of central place populations

observations used in calculating the descriptive parameters, and s_{yx} is the standard error of the estimate, which has the value

$$s_{yx} = \sqrt{((N - 1)(s_y^2 - b^2 s_x^2)) / (N - 2)},$$

in which s_y^2 is the variance of the independent variable. The calculated value t_b is then compared with a standard statistical table showing percentile values of Student's "t-distribution." If the value of t_b is larger or smaller than that range specified in the table for a given number of degrees of freedom (N-2) and a given level of confidence, then the slope of the regression line, b, is assessed to be significantly different from zero.

The interpretation of such a result depends in part on the sign of t_b. Figure II-2B illustrates a situation in which the slope is negative, showing an inverse relationship between city size and spacing.[38] That is, larger cities are more closely spaced than the idealized central place pattern specifies, and the larger a particular city is, the greater is the deviation of its nearest neighbor distance from that which is expected. Figure II-2C, to the contrary, shows a direct relationship between the size of a city and its standardized distance. In this instance, small cities appear

---

[38]It must be remembered that the observed pattern is being compared with the ideal central place arrangement. Therefore, real world distances might appear to not conform to such an inverse type of size-distance relationship.

to be much more clustered than larger cities.[39] In the event that the slope of the regression line is found to be significantly different from zero, further analysis in comparing an observed distribution of settlements with that specified by central place theory is not necessary.

If, however, the value of b in the regression equation is found to be not significantly different from zero, further comparisons with the theoretical distribution may be made. The second analytic model proposed herein involves the y-intercept of the regression equation, a. The test statistic, t_a, is calculated

$$t_a = ((\overline{Y} - a) \cdot \sqrt{N}) \; / \; s_{yx} \qquad (4)$$

where $\overline{Y}$ is either the arithmetic mean of the dependent variable or a magnitude specified by certain underlying assumptions and conditions -- in this instance, central place theory -- which provide an expected value for $\overline{Y}$, and a is the y- intercept (or the value of y_i when x_i is zero). As

---

[39]The horizontal dashed line extending across each of the four graphs in Figure II-2 represents the expected relationship according to Christaller and the compatible modifications of central place theory. Notice that an inverse size-distance relationship is characterized by a regression line which lies <u>below</u> the theoretically expected line. This is so, because as small places reflect less clustering, their average spacing approaches the maximum possible. Increased spacing with increased population, on the other hand, knows no such restraints, and is limited only by the maximum possible distance between two points within the study area.

in the preceding discussion, the calculated value t_a is compared with a table of percentiles of Student's "t-distribution." If the value of t_a is larger or smaller than that range specified by the table, then the spacing of the observed group of cities is significantly different from the expected pattern.[40]

Interpreting the results of this analytic approach requires that the slope of the regression line is not significantly different from zero. Given that this requirement is met, the implications resulting from the test of the y-intercept are as follows. First, consider the simple case where the entire study universe is being analyzed. If the value of a is found to be not significantly smaller than 1.075, then the distribution of settlements is not significantly different from that specified by central place theory. If this is the case, the analysis can end at this point with the conclusion that the settlement pattern resembles that formulated by Christaller.

---

[40]In the comparing of an observed distribution of settlements with that which is specified by central place theory,

$$\overline{Y} = 1.075$$

as discussed earlier. In the event that one is comparing an entire study universe with the defined central place pattern, the y-intercept cannot exceed 1.075. This restriction is not applicable to areas or groups of cities which include less than the entire study universe.

If, on the other hand, the value of the y-intercept is found to be significantly smaller than 1.075, it indicates that the settlements in the study universe do not conform to the central place pattern over the entire area (see Figure II-2A and D). Further analysis is then necessary to properly evaluate the nature of the size-distance relationship of the study area or cities.

Second, consider a more complex situation in which somewhat less than the entire study universe is being analyzed. If the y-intercept is not significantly different from 1.075, it indicates that the settlements within the subregion possess the settlement distributional characteristics which central place theory specifies for the entire study area with respect to the descriptive parameters of the size-distance regression equation. If the calculated value t_a is found to be significantly smaller or larger than 1.075, then it can be observed that the subregion does not possess a settlement distributional pattern which is representative of central place theory for the entire study area. In any event, proper evaluation of the distribution of settlements in any subregion requires further analysis using a third analytic model.

A third test of significance which can be used in the analysis of central place patterns involves the correlation coefficient

$$r = b \cdot s_x / s_y.$$

This descriptive parameter may be used to compute the variable

$$z_r = 1/2 \cdot \ln((1 + r) / (1 - r))$$

which has a nearly normal sampling distribution with mean approximately

$$z_\rho = 1/2 \cdot \ln((1 + \rho) / (1 - \rho))$$

and standard deviation approximately

$$s_z = 1/\sqrt{N-3}.$$

The correlation coefficient, r, can be tested for a significant difference from an expected level of correlation, ρ, with the use of the test statistic

$$Z = (z_r - z_\rho) / s_z. \qquad (5)$$

This analytic model is used only when the slope of the regression line is not significantly different from zero and the y-intercept is significantly different from 1.075 for the entire study universe. A value of one (1) is substituted for ρ, and Z is then calculated.[41] If the calculated value of Z is greater than the normal deviate having a specified percentile rank or confidence level, r is assessed as being

---

[41] If $\rho = 1$, then z_ρ = infinity. It is recommended, therefore, that an arbitrarily high level of correlation be used, say .99 or .999. These two figures yield z_ρ's of 2.645 and 4.0 respectively.

significantly different from ρ, 1. The conclusion which must then be arrived at, of course, is that the distribution of settlements is significantly different from that which would be expected according to central place theory (see Figure II-2A and D). In the instance represented by Figure II-2A, the distribution of settlements would show a low absolute value for the correlation coefficient. Figure II-2D, on the other hand, illustrates a situation in which the distribution of settlements corresponds quite closely to central place theory, but this spatial pattern does not cover the entire area, whether or not it includes the whole study universe.

Influences of Error

Sources of error may be found at several points in determining the parameters for models of this type. In the populations of cities, for example, the sources of data may be subject to bias. As for the distance variable, these data may be affected by measurement errors or bias in computation procedures, such as rounding-off of decimal values. The recognition of data and measurement errors, and of an error function is no simple matter, and it is proper that the effects of such phenomena on the descriptive and analytic models be contemplated at this time.

The influence of such errors on the slope of the regression line, b, must be taken into account in analyzing this descriptive parameter. Walker and Lev state that error in the dependent variable, standardized distance, has little influence on the magnitude of the regression coefficient.[42] Errors in the independent variable, population size, however, tend to reduce the absolute value of the regression coefficient. This is so, since such errors increase the variance of the dependent variable, which has an inverse relationship with b in the equation

$$b = (\Sigma X_i Y_i - (\Sigma X \cdot \Sigma Y) / N) (s_x^2 \cdot (N - 1)).$$

The test statistic, t_b, to the contrary, is influenced by the variance of both the independent and dependent variables. The affects of such errors are direct for the independent variable and inverse for the dependent variable. The influence on t_b, then, is related to the balance of measurement error between the two groups of data. If all such errors are distributed equally between the two variables or are found to lie dominantly in the dependent variable, then the effect is to reduce the chance of finding the slope of the regression line to be significantly different from zero -- the expected value of b as defined by

---

[42]Helen M. Walker and Joseph Lev, Statistical Inference (New York: Holt, Rinehart and Winston, Inc., 1953), p. 306.

central place theory. If, on the other hand, the presence of measurement error lies mainly with the independent variable, population, the effect is to increase the calculated value of t_b, thus counter-acting part of the reduction in b.

The second analytic model, that involving the y-intercept of the regression equation, is also affected by measurement error. The value of this parameter is computed using the slope of the regression line and the arithmetic means of the two variables in the equation

$$a = \overline{Y} - b\overline{X}.$$

Since measurement error has little effect on the arithmetic means, $\overline{X}$ and $\overline{Y}$, the net effect is to reduce the difference between a and $\overline{Y}$.

The test statistic for the y-intercept, t_a, is inversely related to the standard error of the estimate. Therefore, a dominance of measurement error in the independent variable tends to reduce the calculated value of t_a. Both equality in the distribution of such error between the two variables or dominance in dependent variable tend to increase the value of t_a. Thus, as in the case of the regression coefficient, the influence of measurement error is related to the balance of such error between the two variables.

The third analytic model is a test of the value of the correlation coefficient, r. The effect of measurement errors in either or both the independent and dependent variables tend to force the value of r toward zero. The test statistic, Z is also forced toward zero by measurement errors. The width of the tolerance interval, however, is independent of measurement error. Therefore, the presence of such error has the effect of increasing the chance of finding an observed settlement pattern to be significantly different from that associated with central place theory.

Preliminary Evaluation

In the introductory section of the paper, it was pointed out that completed studies of city distribution have been partially unsatisfying. The reasons for this situation were mentioned and discussed. In order to provide for some of these shortcomings, then, the development of a new analytic approach was undertaken.

The first objective of the constructing of these models was to attempt to account for the size-distance relationship of settlement distribution. This was done with the intention of placing a useful combination of techniques before students of central place theory. Also, this theoretical construct provides a foundation of rigorously defined terms and relationships upon which to base such an attempt.

The second objective was to design descriptive models which could produce results capable of undergoing certain tests of significance. This required that the techniques be within the realm of inferential statistics. This group of techniques -- the tests of significance -- allows the researcher to make heretofore impossible statements as to the characteristics of an observed distribution of cities. That is, descriptions of such patterns are no longer restricted to the very general terms uniform, random and clustered.

As a result of attempting to satisfy these two objectives, a new group of models is proposed. The models possess a number of advantages.

(1) They are simple.

That is, the data requirements are small -- only population and standardized distance.

(2) They are inferential.

The parameters of the descriptive models can be subjected to certain tests in order to detect the presence of relationships not readily visible or only suspected to exist.

(3) They are comparative.

Previous discussions have been concerned with comparing observed settlement patterns with that which would be expected according to central place theory. While this has been of primary concern in the development of the descriptive and analytic models, it should also be pointed out that observed parameters from different groups of cities or study areas may be compared using the same techniques.

CHAPTER III

EMPIRICAL DEMONSTRATIONS OF THE MODELS

In the previous two chapters it is argued that present research techniques designed to assess central place theory have been partially unsatisfying. Several works are briefly reviewed and their shortcomings discussed. A pair of descriptive models and three tests are then proposed which are designed to compensate for the aforementioned deficiencies. The validity of these models is established in the following empirical situations.

The Study Universe

The units of study consist of 1054 agglomerated settlements within the state of Michigan. This selection was made for several reasons. First, the distribution of population is not uniform over the entire state. Second, Michigan is characterized by having a diverse economic structure, with a variety of kinds of activities being important in different portions of the state. Third, the peninsularity of the state provides an interesting boundary problem: an actual truncation of developable space exists, a condition which is incompatible with Christaller's assumption of an unbounded plain. Finally, settlements in selected

subregions of the state are used in particular instances in which they possess the needed parameters.

Each of the 1054 cities had a population in 1960 in excess of 100, can be located on a map, and is not situated on an island in the Great Lakes. The models require the following data for each settlement:

(1) name;

(2) population;[43]

(3) latitude and longitude;[44] and

(4) peninsular location.

From the standpoint of developing the proposed models and the works upon which they are based, it is assumed that

---

[43]Three sources of population data were consulted. They are listed according to the priority given their figures.

Bureau of the Census, "Table 8. Populations of All Incorporated Places and Unincorporated Places of 1000 or More, 1960," United States Census of Population, 1960: Michigan: Number of Inhabitants, PC(1)24A (Washington: United States Department of Commerce, 1961), pp. 23-5.

Michigan: Official Highway Map: 1961 (Lansing: Michigan State Highway Department, 1961).

"Michigan: Index of Cities, Towns, Counties, Transportation Lines, Airports, Banks and Post Offices," Commercial Atlas and Marketing Guide: 1962 (Chicago: Rand McNally and Co., 1962), pp. 233-7.

[44]Precise locations are not generally available for most settlements in Michigan. The locational data used in this research was degrees and four place decimals of degrees of latitude and longitude, and was obtained by personal measurement on the following group of maps:

United States Geological Survey, "Michigan" (twenty-seven sheets), National Topographic Map Series: 1:250,000 (Washington: United States Department of the Interior, 1954-1964).

population is a satisfactory index of social and economic complexity of agglomerated settlements.[45] No regard is given to the corporate status of the place, nor to the fact that its location is far from, near to, or, indeed, even within another city.

As to the locations of towns, some generalization is required. Obviously, settlements occupy areas, whereas latitudes and longitudes indicate point locations. At the scale of the maps used, most places are symbolized by small circles which approached point-like character. Larger cities, however, are indicated by area symbols, and usually include several of the more important roads which pass through the place. In these instances, an effort is made to subjectively locate the center of the town and assign the latitude and longitude of that point to the place.

Because the study universe lies on the two main peninsulas of Michigan, it is necessary to take into account on which peninsula a particular town is situated. This is required, because the distances between places must be entirely on land. Therefore, the distances between pairs of points on opposite peninsulas must be by way of Mackinac

---

[45] Brian J. L. Berry and William L. Garrison, "The Functional Bases. . .," loc. cit.

Bridge, the only point where land travel between these two subregions is possible.

Operational Definitions

In order that the fractile diagram may be used to identify an equal size category for each agglomerated settlement, it is necessary that the distribution of populations possess a normal form. The populations of Michigan settlements in 1960 did not exhibit the required distribution of sizes.

Several normalizing transformations were tried, but none proved to be completely successful. That is, none of the transformations altered the data in such a manner so as to fit the necessary number of observations into the ninety percent confidence limits of the fractile diagram test for normality.[46] The transformation

$$x_i = (\log_{10} (\text{population}_i))^{1/2}$$

provided the closest approximation to normality, accounting for 80.11 percent of the observations, while the remaining 19.89 percent of the values lay outside of the specified

[46]When limits associated with a specified degree of confidence are used in a statistical test comparing an observed distribution with a theoretically calculated normal distribution, it is to be expected that at least a percentage of observations equal to the degree of confidence will fall within the relevant tolerance interval (see Appendix A).

confidence interval. Among the 340 transformed observations which fell outside of the acceptance region, less than ten percent were above the median. This indicated that a systematic disturbance existed in the data.

Most of the remaining 307 values which had been rejected belonged to groups of single population sizes, which by the large number of identical magnitudes are suspicious. For example, there are ninety settlements in Michigan in 1960 with populations listed to be 100. Descending in order of observed populations, there are two places with 95 inhabitants, twenty-five with 90, one with 86, six with 85, twenty places with 80 residents, one with 78 and forty-six with 75. Similar patterns exist both above 100 inhabitants and below seventy-five. Such evidence appears to show considerable bias on the part of the data sources toward values ending in zero and five.

A means of circumventing this bias is not readily apparent. As previously noted, however, the higher end of the distribution of transformed populations shows a considerable degree of conformity to a normal distribution. Indeed, if only places with populations greater than one hundred are evaluated, the proportion of the populations which fall outside of the ninety percent acceptance region is only 8.2 percent. Adding places with populations of 100 or less serves only to increase the proportion of the observations which falls into the rejection region to a degree beyond the tolerable limit of ten percent. Hence, only those settlements having populations in excess of 100 and percentile ranks greater than 37.9 are used in the analyses (see Figure III-1, showing ninety-five selected observations). The transformed population data are then substituted into the equation as the independent variable.

The normalized population values are used in the generation of the dependent variables, standardized distances. This is done by first calculating equal size categories about each population figure (see Table III-1). Distances are then computed between each place i and all places j ($j \neq i$) of equal or larger population size.[47] These calculated distances are then standardized by multiplying them by the square root of the density, which is calculated using the number of places found to be equal to or larger than the i^{th} city. Simultaneously, computations are made for the expected real and standardized distances of hexagonal lattices having the number of points just mentioned (see Table III-2).

---

[47] Distances between pairs of points were calculated using the measured latitudes and longitudes of each place and standard spherical trigonometric formulae (see Appendix B).

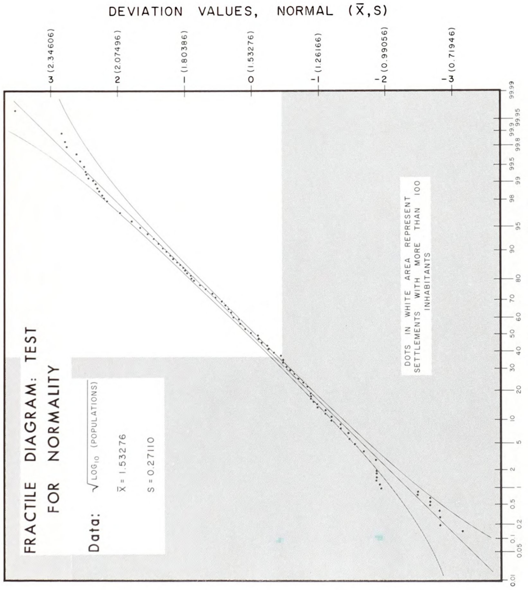

DEVIATION VALUES, NORMAL ($\bar{X}$,S)
3 (2.34606)
2 (2.07496)
1 (1.80386)
0 (1.53276)
-1 (1.26166)
-2 (0.99056)
-3 (0.71946)
FRACTILE DIAGRAM: TEST FOR NORMALITY
Data: $\sqrt{\text{LOG}_{10}}$ (POPULATIONS)
$\bar{X}$ = 1.53276
S = 0.27110
DOTS IN WHITE AREA REPRESENT SETTLEMENTS WITH MORE THAN 100 INHABITANTS
PERCENTILE VALUES, NORMAL (O,I)
0.01
0.05
0.1
0.2
0.5
1
2
5
10
20
30
40
50
60
70
80
90
95
98
99
99.5
99.8
99.9
99.95
99.99

Fig. III-1.

Table III-1. Number of Settlements with Populations Equal to or Larger than Selected Settlements

Selected City	Population	Limits of Equal Population Size Categories: Lower	Limits of Equal Population Size Categories: Upper	Number of Places of a Larger Size than Lower Limit
Dowling	125	114	137	1057
Millet	250	227	275	826
Bath	500	450	556	604
Standale	1000	888	1127	435
Cutlerville	2000	1744	2296	292
Okemos	4000	3388	4730	171
Fair Plain	7998	6540	9803	107
Wayne	16034	12615	20441	70
Southfield	31501	23306	42767	42
Livonia	66702	43884	102202	25
Dearborn	112007	55974	228963	16

Table III-2. Expected Distances Between Selected Settlements[48]

Selected City	Number of Places of a Larger Size than Lower Limit	Density	Expected Real Distance	Expected Standardized Distance
Dowling	1057	1.81565×10^{-2}	7.99 mi.	1.07498
Millet	826	1.41885×10^{-2}	9.02	1.07498
Bath	604	1.03751×10^{-2}	10.56	1.07498
Standale	435	7.47217×10^{-3}	12.44	1.07498
Cutlerville	292	5.01579×10^{-3}	15.18	1.07498
Okemos	171	2.93733×10^{-3}	19.84	1.07498
Fair Plain	107	1.83797×10^{-3}	25.07	1.07498
Wayne	70	1.20241×10^{-3}	31.00	1.07498
Southfield	42	7.21447×10^{-4}	40.02	1.07498
Livonia	25	4.29429×10^{-4}	51.87	1.07498
Dearborn	16	2.74837×10^{-4}	64.86	1.07498

[48] Assuming each is in an hexagonal lattice having as many points as there are places of a larger population than the lower equal size limit

Tests using the fractile diagram show that the distribution of standardized distances, like that of the populations, does not conform to normality. More than a dozen normalizing transformation functions were applied to this data, but even that which provided the best results could arrange only 61.00 percent of the observed values between the specified ninety percent confidence limits centered about a theoretical normal distribution. The transformation function

$$r_i^{1/3}$$

in which r_i is the observed standardized distance, gives the best visual appearance on the fractile diagram as well as the best calculated comparability to values expected in a normal distribution having the same mean and variance (see Figure III-2, showing ninety-five selected observations). This set of transformed values is then substituted into the regression equation as the dependent variable.

Applications and Tests of the Models

Arguments have been offered in order to form a basis for a new approach to the analysis of the spatial distribution of cities. Centered about the relationships presented or inferred within the contexts of central place theory and its recent expansions, this group of techniques allows the recognition of patterns of settlement spacing heretofore

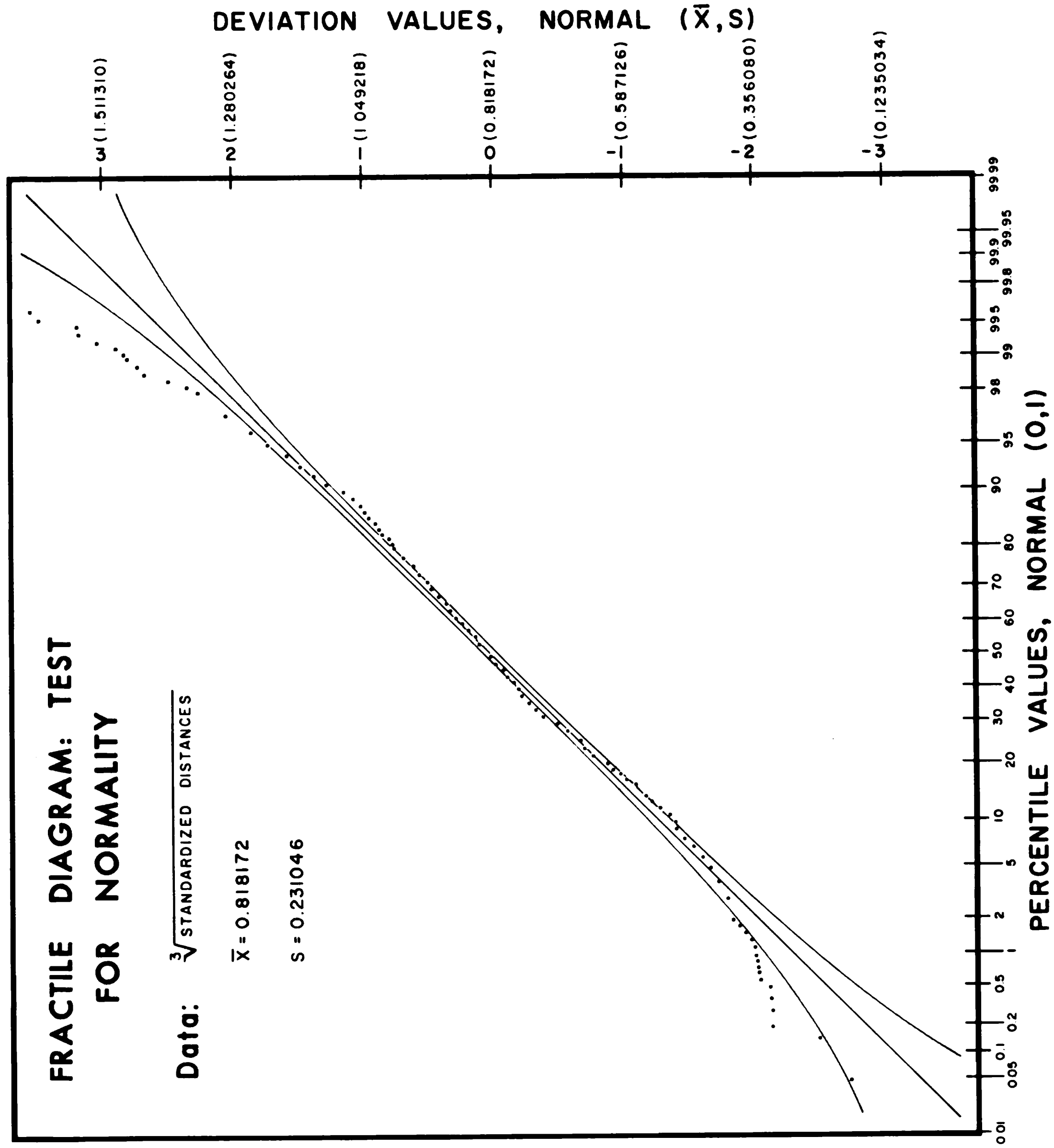

FRACTILE DIAGRAM: TEST FOR NORMALITY
Data: $\sqrt[3]{\text{STANDARDIZED DISTANCES}}$
$\bar{X}$ = 0.818172
S = 0.231046
DEVIATION VALUES, NORMAL ($\bar{X}$,S)
3 (1.511310)
2 (1.280264)
1 (1.049218)
0 (0.818172)
-1 (0.587126)
-2 (0.356080)
-3 (0.1235034)
PERCENTILE VALUES, NORMAL (O,I)
0.01
0.05
0.1
0.2
0.5
1
2
5
10
20
30
40
50
60
70
80
90
95
98
99
99.5
99.8
99.9
99.95
99.99

Fig. III-2.

undiscernable with such simple models. The following discussions relate the results of applications of the models, and are based upon the aforementioned arguments and relationships, and upon the operational definitions of city size and nearest neighbor distance as are specified in the previous section of this paper.

The First Test of Significance

The first analytic approach involves the regression coefficient of a simple regression relationship, equation (1). Based upon the operational definitions of settlement population and standardized nearest neighbor distance, this equation has as a more specific form

$$y_i^{1/3} = a + b(\log_{10}x_i)^{1/2},$$

in which x_i is the population size of the i^{th} place and y_i is the standardized distance from place i to its nearest neighboring place j of equal or larger size. The analytic tool concerned with the regression coefficient is the test statistic in which t_b is calculated, equation (3).

The first application of this test statistic involves the regression coefficient for the entire study universe (see Figure III-3). In this instance, all of the settlements in Michigan with populations of more than 100 are used in calculating a simple regression equation. This equation has the form

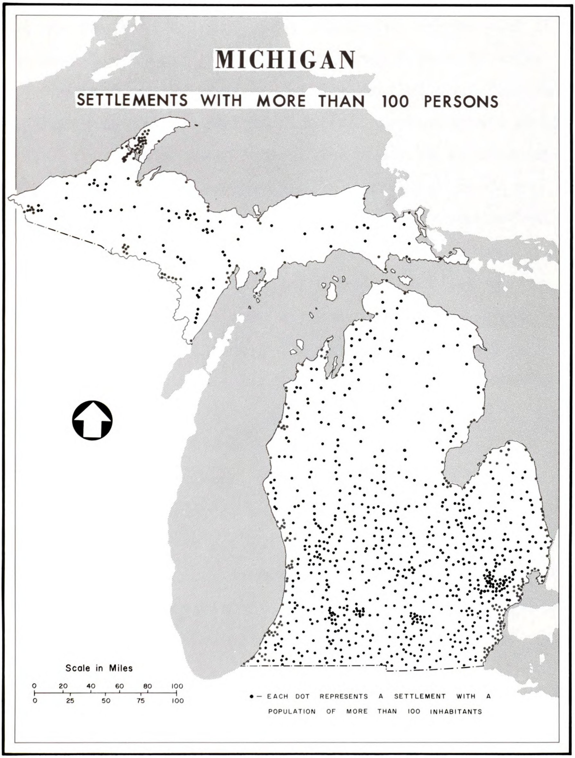

MICHIGAN
SETTLEMENTS WITH MORE THAN 100 PERSONS
Scale in Miles
0 20 40 60 80 100
0 25 50 75 100
• — EACH DOT REPRESENTS A SETTLEMENT WITH A POPULATION OF MORE THAN 100 INHABITANTS

Fig. III-3.

$$y_i^{1/3} = 1.23646 - (0.24668 \cdot (\log_{10}x_i)^{1/2}),$$

and the regression coefficient, -0.24668, may be used in the analytic model to calculate t_b (see Figure III-4A).

As part of the analysis of the spatial distribution of agglomerated settlements in Michigan, an hypothesis can be stated concerning the nature of the observed size-distance relationship and its comparison to that which is to be expected according to central place theory. Subjectively, this hypothesis could read:

> The size-distance relationship for Michigan settlements larger than 100, as indicated by the regression coefficient, is not significantly different from that which is dictated by central place theory.

Mathematically, the same hypothesis may be stated:

$$b = 0.0.$$

The null hypothesis, that statement which is to be accepted if the hypothesis is found to be not true, is

$$b \neq 0.0.$$

If the null hypothesis is accepted, it must be concluded that there is a bias in the spacing of Michigan cities which is related to town population and which does not conform with central place theory.

The calculated value of t_b includes the specific value

$$t_b = ((-0.2467 - 0.0) \cdot 0.1882 \cdot \sqrt{1054 - 1}) / 0.2263$$

which results in

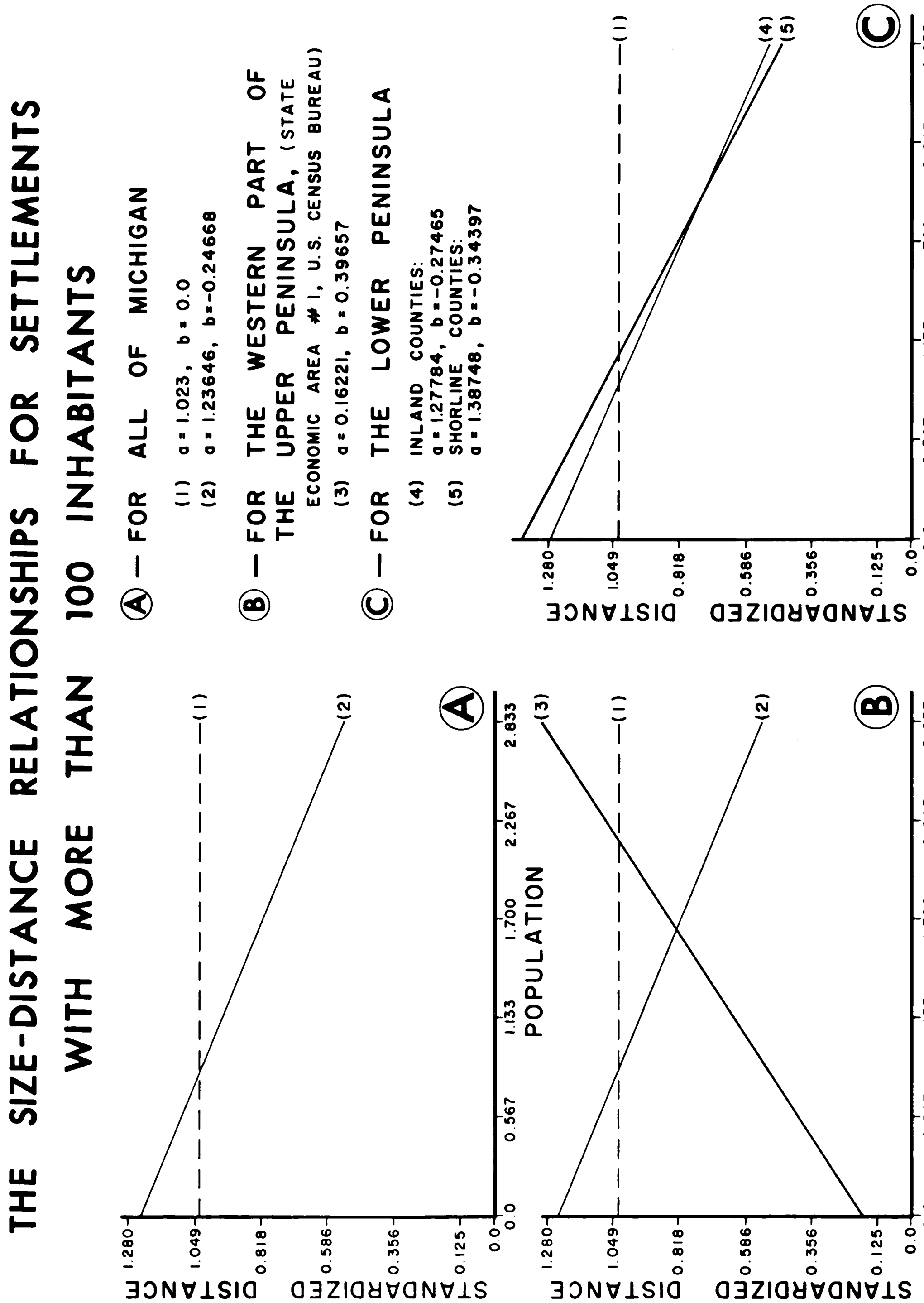

THE SIZE-DISTANCE RELATIONSHIPS FOR SETTLEMENTS WITH MORE THAN 100 INHABITANTS
A — FOR ALL OF MICHIGAN
(1) a = 1.023, b = 0.0
(2) a = 1.23646, b = -0.24668
B — FOR THE WESTERN PART OF THE UPPER PENINSULA, (STATE ECONOMIC AREA #1, U.S. CENSUS BUREAU)
(3) a = 0.16221, b = 0.39657
C — FOR THE LOWER PENINSULA
(4) INLAND COUNTIES: a = 1.27784, b = -0.27465
(5) SHORLINE COUNTIES: a = 1.38748, b = -0.34397
STANDARDIZED DISTANCE
1.280
1.049
0.818
0.586
0.356
0.125
0.0
0.567
1.133
1.700
2.267
2.833
POPULATION
(1)
(2)
(3)
(4)
(5)
A
B
C

Fig. III-4.

$$t_b = -6.5654.$$

The calculated value of t_b is then compared with a predefined entry in a table of the percentiles of Student's t-distribution. If t_b does not fall within the range of zero plus or minus $t_{(\%,df)}$, then b is found to be significantly different from zero at a prespecified level of confidence.[49]

A critical value of t from such a table is

$$t_{(0.5,1052)} = 1.96.$$

Since the range 0.0 ± 1.96 does not include the calculated value of t_b, it must be concluded that t_b is significantly different from zero and, therefore, that

$$b \neq 0.0$$

at the .05 level of confidence.

Thus, it is reasonable to state that among Michigan cities having more than 100 inhabitants there exists a bias in the size distance relationship. This bias appears to be inconsistent with the pattern which would be generated if the rules and assumptions of central place theory were applied to the study universe. Noting the negative slope of the regression line, it can also be observed that larger cities in Michigan tend to be more clustered with respect

---

[49]The subscripts of the value of t from a table of percentiles, °/o and df, indicate the proportionate size of the rejection region and the number of degrees of freedom allowed.

to their nearest neighbors of equal or larger population-size than do smaller settlements in the state.

A second test of significance involves the selection of a "sample" region out of the study area and the comparing of its settlement pattern with that of the entire universe. The sample region consists of the counties forming State Economic Area 1, the western portion of Michigan's Upper Peninsula (see Figure III-5). This area was chosen because of the high proportion of the region's employment in mining activities (eighteen percent).[50] This characteristic, of course, renders the economic structure different from that of the state as a whole, and gives the area qualities which are incompatible with central place theory.

In this test, since the employment structure of the area is quite different from that of the state as a unit, a reasonable hypothesis might contend that the settlement pattern will also differ. It could be stated:

> The relationship between settlement size and spacing in Michigan's S.E.A.-1, as indicated by the regression coefficient, is significantly different from that of the state as a whole.

---

[50]Donald J. Bogue and Calvin L. Beal, "Michigan: Upper Peninsula: Western Area," Economic Areas of the United States (New York: The Free Press of Glencoe, Inc., 1961), pp. 756-8.

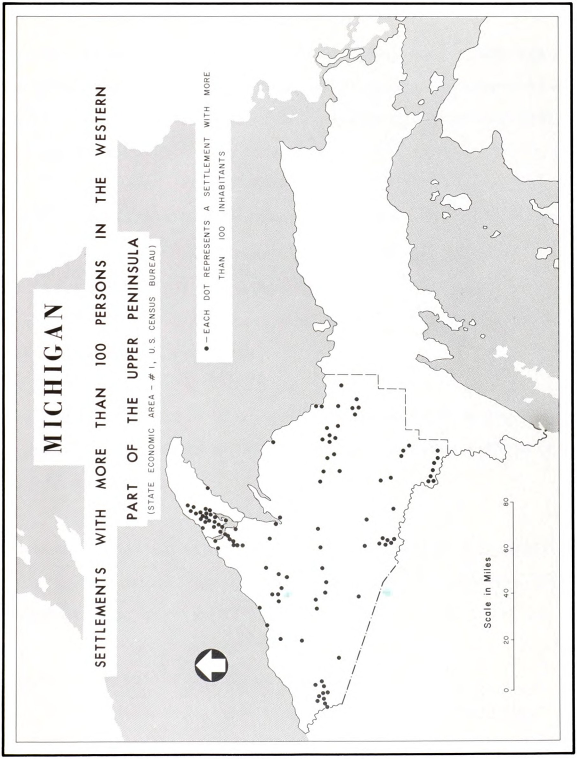

MICHIGAN
SETTLEMENTS WITH MORE THAN 100 PERSONS IN THE WESTERN PART OF THE UPPER PENINSULA
(STATE ECONOMIC AREA - # I, U. S. CENSUS BUREAU)
• — EACH DOT REPRESENTS A SETTLEMENT WITH MORE THAN 100 INHABITANTS
Scale in Miles
0
20
40
60
80

Fig. III-5.

Or, more mathematically,

$$b \neq -0.2467.$$

The null hypothesis, by definition, then, must claim that the size-distance relationship, as shown by the regression coefficient, is not significantly different from the state as a unit, and, thus

$$b = -0.2467.$$

The descriptive model for this example has the form

$$y_i^{1/3} = 0.1622 + (0.3966 \cdot (\log_{10} x_i)^{1/2})$$

(see Figure III-4B), and the significance test has the form

$$t_b = ((0.3966 + 0.2467) \cdot 0.1483 \cdot \sqrt{110 - 1}) \,/\, 0.2955,$$

which generates a value of

$$t_b = 3.3821.$$

This calculated value of t_b is then compared with that range of permissible values as specified by the table of the t-distribution, in which

$$t_{(.05,108)} = 1.98.$$

Since the range of 0.0 ± 1.98 fails to include the calculated value of t_b, it can be concluded that t_b is significantly different from zero and, therefore, that

$$b \neq -0.2467$$

at the .05 level of confidence.

This result leads one to accept the hypothesis that the size-distance relationship for settlements in the western

portion of the Upper Peninsula of Michigan, as shown by the regression coefficient, is significantly different from that of the state as a whole. It can be concluded, then, that this sample of towns, based on an arbitrarily delimited subregion, does not reflect the settlement pattern found to be characteristic of the study universe. A small amount of additional information may be acquired by noting that the sign of the regression coefficient is positive, indicating that a greater degree of relative clustering exists among smaller centers than among larger ones.[51]

A third illustration of the first test of significance involves the comparison of two "sample" regions. The regions used in this example consist of those counties in Michigan's Lower Peninsula which border on a shoreline of one of the Great Lakes, and of the remaining or inland counties of the same peninsula (see Figures III-6 and 7). These subregions have been selected because of their obvious association with a fundamental geographic puzzle -- the boundary problem. One area is surrounded on all sides by space which is capable

---

[51]A subsequent calculation of t_b in which the regression coefficient 0.3966 is compared with zero, that which would be expected according to central place theory, produces a value of 2.089. This recomputed t_b also lies outside of the range 0.0 ± 1.98, thus giving substance to Christaller's view that mining towns cannot be expected to conform to central place theory.

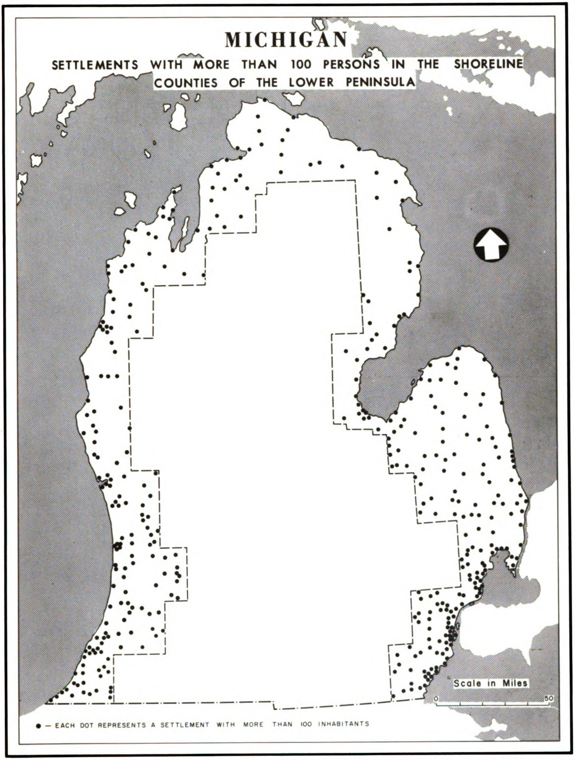

MICHIGAN
SETTLEMENTS WITH MORE THAN 100 PERSONS IN THE SHORELINE COUNTIES OF THE LOWER PENINSULA
Scale in Miles
0
50
● — EACH DOT REPRESENTS A SETTLEMENT WITH MORE THAN 100 INHABITANTS

Fig. III-6.

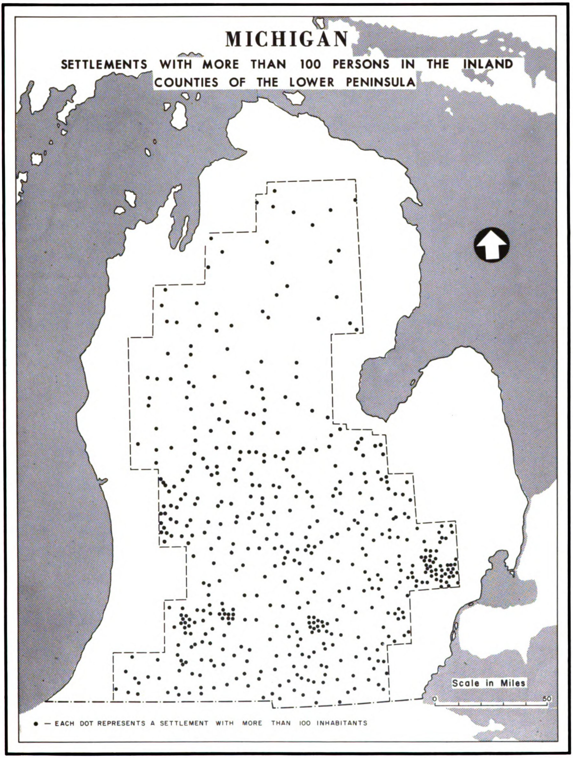

MICHIGAN
SETTLEMENTS WITH MORE THAN 100 PERSONS IN THE INLAND COUNTIES OF THE LOWER PENINSULA
Scale in Miles
0
50
• — EACH DOT REPRESENTS A SETTLEMENT WITH MORE THAN 100 INHABITANTS

Fig. III-7.

of being settled in much the same fashion as the subregion itself, whereas the second area posses a boundary beyond which further settlement is impossible. Also, the presence of a boundary to the area of potential development is not consistent with central place theory.

In this third example, it is suspected that the presence of a shoreline serves to cause a recognizable disturbance of the size-distance relationship of nearby settlements. As a result the regression coefficient associated with the settlement pattern of shoreline counties is expected to be significantly different from that of more inland counties. Such an hypothesis may be stated:

> The relationship between city population and distance to the nearest neighbor of equal or larger size for settlements in the shoreline counties, as indicated by the regression coefficient, is significantly different from that of inland county settlements.

Or, in the terms of an inequality,

$$b_s \neq b_i,$$

where b_s is the regression coefficient for the settlements of the shoreline counties, and b_i is the regression coefficient for the inland places. The null hypothesis,

$$b_s = b_i,$$

must lead to the conclusion that the size-distance relationship

of the shoreline counties, as indicated by the relevant regression coefficients, is not significantly different from that of the inland counties (see Figure III-4C).

Calculating the required parameters for this test of significance produces two descriptive models. For the settlements in the inland counties, the simple regression equation has the form

$$y_i^{1/3} = 1.2778 - (0.2741 \cdot (\log_{10} x_i)^{1/2}).$$

For those in the shoreline counties, the descriptive model has the form

$$y_i^{1/3} = 1.3875 - (0.3440 \cdot (\log_{10} x_i)^{1/2}).$$

The two regression coefficients are then used in the test of significance.

The analytic approach, in this instance, has the form

$$t_b = ((b_s - b_i) \cdot s_{x_s} \cdot \sqrt{N_s - 1}) \;/\; s_{yx_s},$$

in which the subscript s indicates values associated with the shoreline counties, and the subscript i designates the regression coefficient of the inland counties. These variables are used, because the object is to find if b_s is significantly different from b_i. If the object of the test were to find out if b_i is significantly different from b_s, the relevant subscripts would be reversed.

In calculating t_b, the test of significance incorporates the specific values:

$$t_b = ((-0.3440 - 0.2741) \cdot 0.2091 \cdot \sqrt{363 - 1}) \;/\; 0.2227$$

which generate the value

$$t_b = -1.2456$$

This calculated value of t_b is then compared with a critical figure in a table of percentiles of the t-distributions,

$$t_{(.05,361)} = 1.97.$$

Since the calculated value of t_b, -1.2456, falls within the range of 0.0 ± 1.97. t_b is found to be not significantly different from zero and, therefore, b_s is not significantly different from b_i at the .05 level of significance.

Thus, the size-distance relationship for the settlements in the shoreline counties of Michigan's Lower Peninsula is not significantly different from that of the inland counties. A reasonable conclusion, then, is to accept the null hypothesis which proposes that the rates of change in the dependent variables for a unit change in the independent variables are found to be not significantly different from one another. In this instance, both regression coefficients indicate that relative clustering is greater among larger cities than among smaller ones.

The Second Test of Significance

The second analytic approach is a statistical test of significance involving the y-intercept of the same descriptive model, the simple regression equation,

$$y_i^{1/3} = a + (b \cdot (\log_{10} x_i)^{1/2}).$$

The test statistic for this descriptive parameter is equation (3), in which $\overline{Y}$ is either the arithmetic mean of the dependent variable or some previously specified figure -- an expected value according to a set of predefined conditions. The calculated value t_a is compared with a table of percentiles of Student's t-distribution from which is selected a suitable relevant figure, in order to measure the comparability of $\overline{Y}$ and a.

The testing of the y-intercept should be performed only after it has been ascertained that the slope (or slopes) of the regression line(s) is (or are) not significantly different from zero, since as b nears zero, the y-intercept, a, becomes more representative of the average spacing characteristics of the settlement pattern being analyzed. Thus, differences between a group of data and a hypothetical situation, or between two groups of data may be discerned, even if the slopes of the relevant regression lines are not significantly different from zero.

The second test of significance compares the spacing characteristics of the settlements in Michigan's State Economic Area number six with that which would be expected if the rules, assumptions and relationships of central place theory applied to the entire state (see Figure III-8). The descriptive, simple regression equation upon which the analytic model is based contains the following specific parameters (see Figure III-9D):

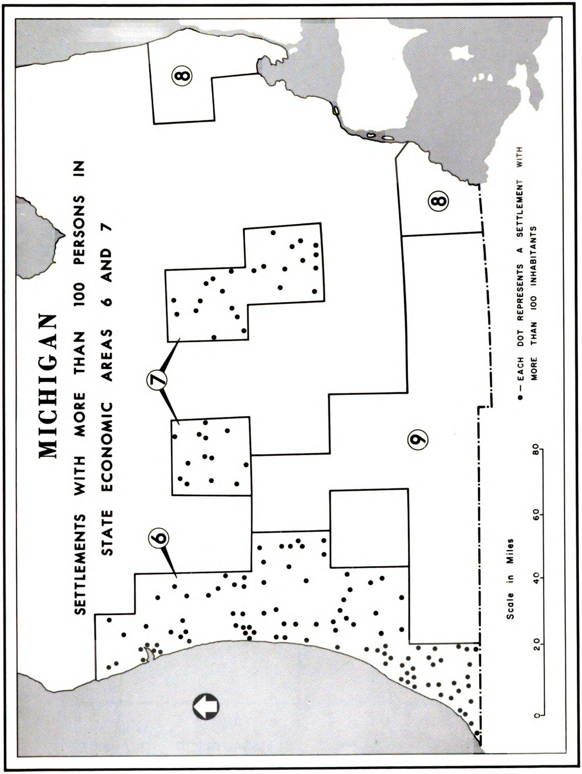

MICHIGAN
SETTLEMENTS WITH MORE THAN 100 PERSONS IN
STATE ECONOMIC AREAS 6 AND 7
6
7
8
8
9
Scale in Miles
0 20 40 60 80
• — EACH DOT REPRESENTS A SETTLEMENT WITH
MORE THAN 100 INHABITANTS

Fig. III-8.

THE SIZE-DISTANCE RELATIONSHIPS FOR SETTLEMENTS WITH MORE THAN 100 INHABITANTS

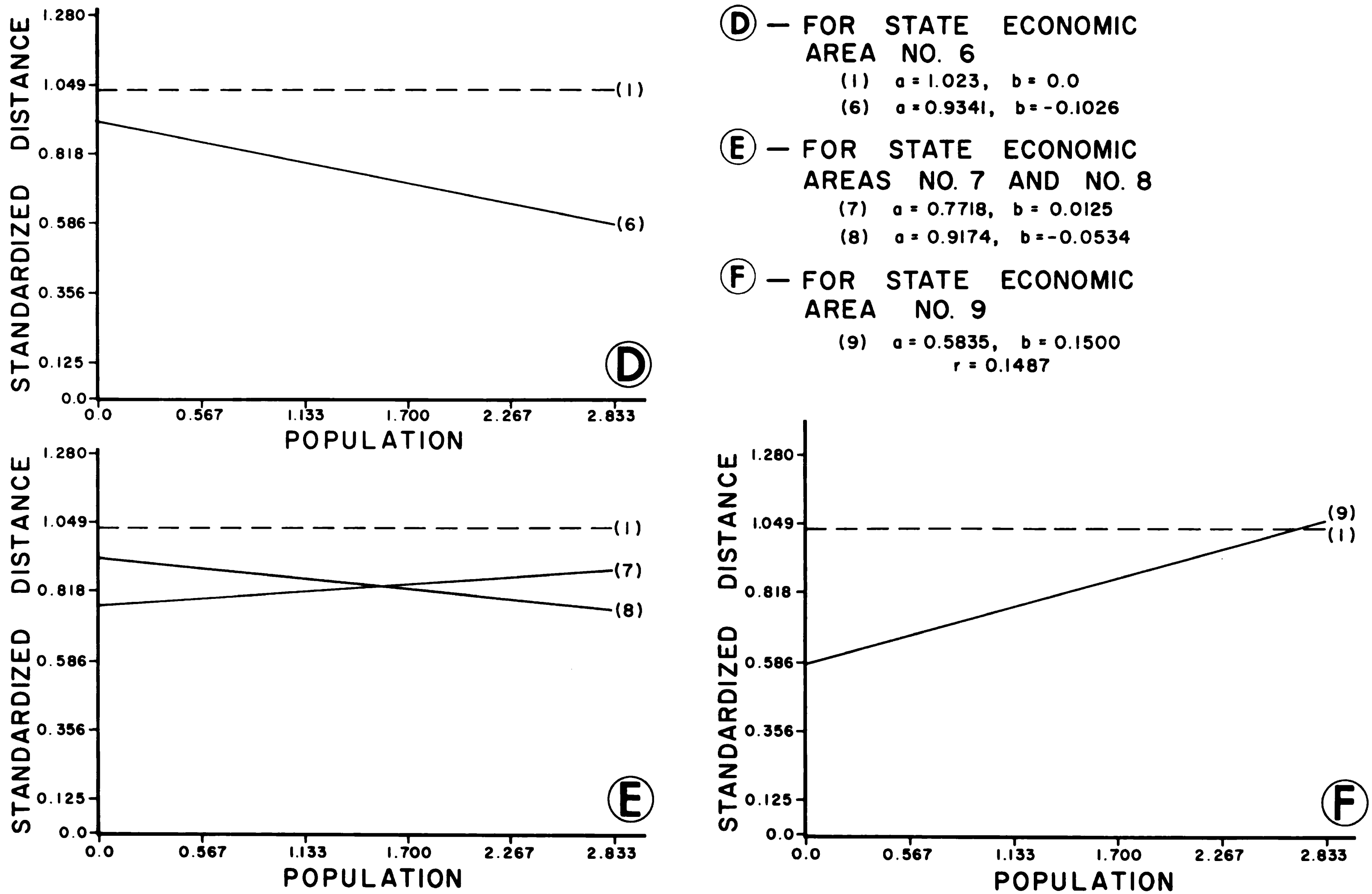


Fig. III-9.

$$y_i^{1/3} = 0.9341 - (0.1026 \cdot (\log_{10} x_i)^{1/2}).$$[52]

Based upon this model and the information that b is not significantly different from zero, a suitable hypothesis requiring the use of the second test might read:

> The spacing characteristics of the settlements in State Economic Area six, as indicated by the y-intercept, are significantly different from those that would be expected if central place theory applied to the state as a whole.

Or, in mathematical terms,

$$a \neq (1.075)^{1/3}.$$

The null hypothesis holds that the spacing characteristics of the settlements in S.E.A.-6 do resemble those which would be expected if central place theory applied to the entire state. The null hypothesis, then, claims that

$$a = (1.075)^{1/3}.$$

This test of significance has the specific form

$$t_a = ((0.9341 - 1.0244) \cdot \sqrt{105}) \ / \ 0.1814.$$

These figures produce a calculated value

$$t_a = 28.1153,$$

---

[52]The regression coefficient, -0.1026, was found to be not significantly different from zero. This was so, because the calculated value of t_b, -1.2679, fell within the range 0.0 ± 1.98, which was based on a $t_{(.05,103)}$ of 1.98.

which, when compared with a critical figure from a table of percentiles of Student's t-distribution,

$$t_{(.05,103)} = 1.98,$$

demonstrates that t_a is significantly different from zero. That is, t_a lies outside of the defined rejection region, 0.0 ± 1.98. This result leads to the conclusion that the observed value of a is significantly different from 1.0244, that which would be expected if an idealized central place pattern existed among Michigan settlements with more than 100 inhabitants.

Interpretation of the result of this comparison allows the acquisition of only a small amount of additional information. Since the calculated value of t_a possesses a minus sign, it indicates that the value of a is smaller than that of $\overline{Y}$. Based on this and the fact that a was found to be significantly different from $\overline{Y}$, it can be observed that the settlements in State Economic Area six are significantly more closely spaced than they would be under the uniform distributional characteristics of central place theory.

Another illustration of the second analytic model can be made by comparing the spacing characteristics of settlements in two subregions of the state. The areas selected for this analyze are State Economic Areas eight and nine, both of which are found in the southern portion of the Lower Peninsula (see Figure III-10). In both of

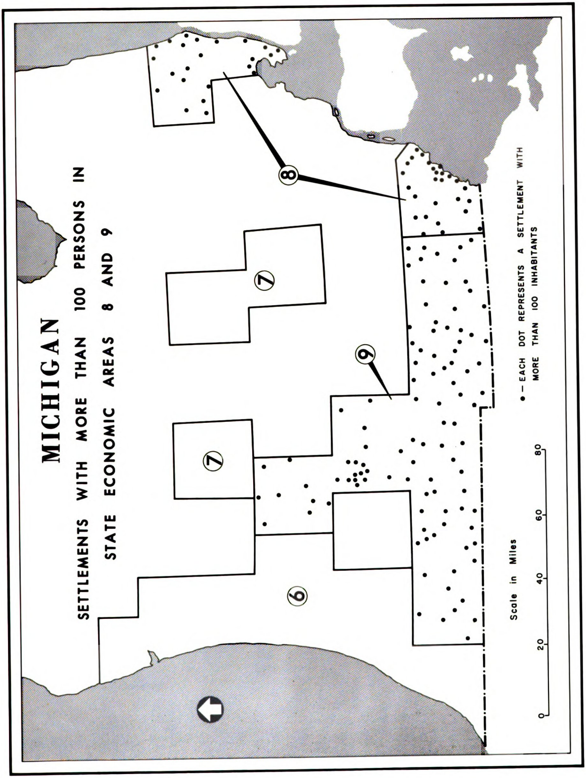

MICHIGAN
SETTLEMENTS WITH MORE THAN 100 PERSONS IN
STATE ECONOMIC AREAS 8 AND 9
6
7
7
8
9
• — EACH DOT REPRESENTS A SETTLEMENT WITH
MORE THAN 100 INHABITANTS
Scale in Miles
0
20
40
60
80

Fig. III-10.

these subregions, the size-distance relationship shows no significant bias with respect to differential clustering from one end of the array of population sizes to the other.

State Economic Area nine is described by Bogue and Beal as lying "near the northern border of the Corn Belt."[53] Agriculture accounts for a large portion of this area's economic activity, although a variety of manufacturing establishments are found in the towns included in S.E.A.-9. The other subregion, State Economic Area eight is described as being "submetropolitan," both socially and economically.[54] Flanking the Detroit urban complex, the economics of this subregion are closely linked with those of the urban area, both from the point of view of there being a large number of manufacturing firms located within S.E.A.-8, and from the standpoint that the nearby Detroit area serves as an employment source for many residents of this subregion. Thus, State Economic Area nine appears to possess certain of the qualities conducive to the development of a settlement pattern approaching that which was theorized

---

[53]Donald J. Bogue and Calvin L. Beal, "Michigan: Southern Michigan: Eastern Area," and ". . . Western Area," loc. cit., pp. 769-71.

[54]Donald J. Bogue and Calvin L. Beal, "Michigan: Southeast Michigan Area," loc. cit., p. 767-9.

by Christaller, whereas the S.E.A.-8 is a suburban area near a large metropolitan complex -- one type of situation which central place theory fails to consider.

State Economic Area nine, then, might be expected to conform more closely with central place theory than S.E.A.-8. In this test, the spacing characteristics of the agricultural area are accepted as the expected situation, and those of the suburban areas are suspected of being significantly different from the theoretical pattern. Therefore, a reasonable hypothesis to be tested would state:

> The spacing of settlements in State Economic Area eight is significantly smaller, as indicated by the y-intercepts, than that of the settlements in State Economic Area nine.

That is, since the regression coefficient for both subregions are not significantly different from zero, it is to be expected that

$$a_8 < a_9,$$

where the subscripts 8 and 9 indicate the respective economic areas. Such a result would be compatible with Christaller's logic, whereas the null hypothesis,

$$a_8 \geq a_9$$

would not. Subjectively, the null hypothesis must claim that the spacing of the suburban settlement pattern found

in S.E.A.-8 is not significantly more clustered than that of S.E.A.-9.

The descriptive models for the two subregions are as follows (see Figure III-9E): for S.E.A.-8

$$y_i^{1/3} = 0.7718 - (0.0125 \cdot (\log_{10} x_i)^{1/2}),$$

and for S.E.A.-9

$$y_i^{1/3} = 0.9174 - (0.0534 \cdot (\log_{10} x_i)^{1/2}).[55]$$

The analytic approach with which a_8 is tested for a significant difference from a_9 is

$$t_a + ((0.7719 - 0.9174) \cdot \sqrt{57}) / 0.1741.$$

The calculated value of t_a is -6.3136. This figure does not fall within the rejection region, -1.67 to plus infinity. Therefore, it is reasonable to accept the hypothesis that

$$a_8 < a_9,$$

and that settlements in S.E.A.-8 are significantly more closely spaced than those found in S.E.A.-9.

The Third Test of Significance

The third test of significance involves the third descriptive parameter, the correlation coefficient, r.

---

[55] The calculated values of t_b are for S.E.A.-8, 0.0845, and for S.E.A.-9, -0.8491. Both of these figures were found to be not significantly different from zero at the .05 level of significance. Thus, neither regression coefficient is found to be significantly different from zero.

This is accomplished by comparing an observed value of r with that value which might be expected according to some preconceived notions. If the difference between the observed and specified values is greater than a given limit, the observed value is adjudged to be significantly different from the predefined one.

An example of this test of significance can be shown using State Economic Area seven (see Figure III-8). The settlement pattern of this area can be described by the simple regression equation

$$y_i^{1/3} = (0.5835 + (0.1500 \cdot (\log_{10}x_i)^{1/2}),$$

which indicates that there is no significant bias in the size-distance relationship (t_b = 1.0657, which falls within the range of 0.0 $\pm$ 2.01), and that the spacing of the settlements in the area is significantly smaller than that which would be expected if the idealized central place pattern applied to the state as a whole (t_a = -157.0992, which falls outside the range of 0.0 $\pm$ 2.01). Since no significant bias in the size-distance relationship is found in this area, the next point of interest is how well the distribution of standardized distances conforms to that of the central place pattern (see Figure III-9F).

If central place theory is applicable within an area, it is to be expected that the correlation of standardized

distance with population should approach unity. In the instance of S.E.A.-7, the descriptive model has the specific form

$$r = (0.1500 \cdot 0.1419) / 0.1431,$$

which generates a value of 0.1487 for the correlation coefficient. Since this value appears to be close to zero, a reasonable hypothesis would read:

> The degree of correspondence between population and standardized distance between a given place and its nearest neighbor of equal or larger size, as indicated by the correlation coefficient, is significantly different from that which would be expected according to central place theory.

Or, in mathematical terms

$$r \ll 1.$$

The null hypothesis, on the other hand, claims that the correlation coefficient is not significantly smaller than one, and that the observed pattern is comparable to that specified by central place theory.

The test statistic has the specific form

$$Z = (0.1497 - 2.6492) / 0.1414,$$

in which

$$z_r = 0.1497, \; r = 0.1487$$

$$z_\rho = 2.6492, \; \rho = 0.9900.$$

The calculated value of Z is -17.6742, which falls outside of the rejection region -1.645 to plus infinity at the .05 level of significance. Thus, the hypothesis that

$$r < 1$$

is accepted. And, the contention that the observed pattern is not significantly different from central place theory is rejected.

Additional interpretation of the results of this test is difficult. On the basis of what has been shown, the only statement which can be made is that while no significant systematic bias has been demonstrated in the settlement pattern of S.E.A.-7, there has also been no success in demonstrating the presence of spatial order the likes of which is defined by central place theory.

CHAPTER IV

CONCLUSION

Earlier discussions have been concerned with the recognition and at least partial solution of certain problems associated with the analysis of central place patterns. In order to arrive at these solutions, it was necessary to apply two descriptive models to two variables which are relevant to the spatial dynamics of settlement distribution -- population size and standardized distance to the nearest neighbor of equal or larger size. The descriptive parameters of these two models are tested for significant differences from values which are predefined according to the rules and assumptions of either central place theory or some other hypothesized situation. Evaluation of the character of an observed settlement pattern can then be based on the results of the analytic models.

Evaluation of the Empirical Applications of the Test of Significance

The study universe, consisting of 1054 agglomerated settlements in Michigan with more than one hundred inhabitants, provided the two types of data necessary for the

applying of the proposed models and tests to a variety of empirical situations. The value for each observation in the independent variable -- population size -- was acquired from already published sources, whereas the value for each observation in the dependent variable -- standardized distance from each given place to its nearest neighboring center of equal or larger size -- was calculated using measured latitude and longitude point locations. Each of the demonstrations relied upon all or part of the study universe, and discussed a problem which has some degree of relevance to the study of settlement distribution in general. That is, the problems dealt with are not unique to the state of Michigan, but are of widespread interest to students of settlement geography from both the regional and systematic standpoints.

Essentially, two types of applications can be made of the tests of significance. One is the evaluation of an observed pattern of settlement distribution with respect to a spatial arrangement specified by some conceptual framework. In this instance, each of the three descriptive parameters is compared with an expected value which can be defined according to the rules and assumptions of central place theory.

The first analytic model involved the regression coefficient. The arguments in the preceding discussions indicate

that the slope of the regression line according to central place theory should be zero. The observed value of b for the entire study universe was found to be significantly smaller than zero, indicating that larger cities in Michigan are relatively more closely spaced than are smaller places. That is, the standardized distances for larger places show greater deviations from the theoretically expected value (or the negative side) than do the standardized distances for smaller settlements. The observed value of b for the settlements located in State Economic Area one was found to be significantly larger than zero, indicating that larger cities in the western part of the Upper Peninsula are relatively less closely spaced than are smaller settlements. Thus, these two analyses revealed biases in the size-distance relationships of the entire study universe and the specified subregion which do not conform to central place theory.

The second analytic approach utilizes the y-intercept. Based upon the materials cited, it was realized that according to central place theory, the expected value of this figure is 1.075. In the one demonstration of the test of significance involving this value, it was found that the observed value for a for a selected subregion, State Economic Area six, is significantly smaller than that

which is to be expected according to Christaller's theoretical formulation. Having already ascertained that the value of the regression coefficient for S.E.A.-6 is not significantly different from zero, such a result was taken to indicate that the settlements in the subregion are spaced more closely than would be the case if central place theory applied to the entire state.

The third test, like the second, requires that the slope of the regression line does not differ significantly from zero. If this is the case, and should the settlements under study conform to a central place pattern covering all or part of the study area, the value of the correlation coefficient can be expected to approach unity. Results of the use of the analytic model involving r showed that the observed value of this parameter for S.E.A.-7 is significantly smaller than unity. This indicates that the independent variable is at best a poor predictor for the dependent variable -- a situation quite different from that which is to be expected according to a deterministic model such as central place theory.

A second application of the tests of significance is the evaluation of two observed patterns of settlement distribution, one with respect to the other. In this type of approach, the necessary descriptive parameters for both settlement distributions must be calculated. One of the

computed values is then compared with the other, in the same fashion as though it were being tested against a theoretically defined value.

The first test of significance, that involving the regression coefficients, was used in two such comparisons of Michigan settlement patterns. In the first of the two, it was found that the size-distance relationship, the amount of change in settlement spacing per unit change in settlement population, for State Economic Area one is significantly different from that of the state as a whole. In the second example, it was found that the size distance relationship for settlements in the shoreline counties of the Lower Peninsula is not significantly different from that of the settlements in the inland counties. Thus, the size-distance relationship for any arbitrary region can be compared with that of any other arbitrary region, and the presence or absence of significant differences can provide information as to the characteristics of the settlement patterns of the two areas.

The second test can indicate relative differences in the spacing of settlements in different regions. In the related analysis involving Michigan, the spacing of settlements in S.E.A.-8 was found to be significantly smaller than that of the settlements in S.E.A.-9. Thus, the distances between places in the metropolitan-suburban

subregion were found to be significantly less than those associated with settlements located in the agricultural area.

Evaluation of the Analytic Models

The purpose of the research involved in preparing this paper is the development of a new approach to the analysis of observed settlement patterns with respect to the deterministic model proposed by Christaller. The accomplishment of this objective requires the recognition of certain characteristics of the idealized central place pattern of settlement distribution. Upon their recognition, models can be developed with which observed values for these characteristics can be compared with the relevant theoretically defined figures.

Three characteristics inherent to the spatial pattern of settlement distribution as conceptualized by Christaller are identified. One of these describes the relationship between the population size and nearest neighbor distance for a group of settlements. The second is an indication of settlement spacing, given that the size-distance relationship conforms to that specified by central place theory. The third characteristic which was recognized was that a close degree of correspondence must exist between city size and distance to the nearest place of equal or

larger size. A numeric value for each of these three descriptive parameters can be calculated with the use of two descriptive models.

An observed value for each of these characteristics can be analyzed according to inferential statistical methods. These analyses provide bases for the making of heretofore impossible statements concerning the nature of settlement distributions. As a consequence, verbal terminology for the description of settlement patterns is no longer restricted to the terms "more uniform than random," "random," and "more clustered than random."

An expanded and somewhat more valuable terminology is, therefore, proposed. In it are included the terms "uniform," "significantly different from uniform," "random," "significantly different from random," "positive differential clustering" (as in the instance of State Economic Area one), and, finally, "negative differential clustering" (as in the instance of the state as a whole). In addition, a basis for a comparative phraseology has also been formed. It is now possible to note that places in one area are "relatively more closely spaced" or "relatively less closely spaced" than in a second area, or that the degree of differential clustering in one region is "significantly different from" that found in a second area.

Essentially, what is discussed in this paper is a pair of general models with which the distance -- population size relationship of settlement distribution may be described. These models, simple as they are, provide the researcher with three parameters that indicate the nature of this important central place relationship for any observed group of cities. In addition, tests of these parameters are capable of providing information as to the following:

(1) conformity to central place theory over the entire study area;

(2) conformity to central place theory over less than the entire study area;

(3) a random distribution of cities of all sizes within the study area;

(4) differential clustering among cities of various population sizes.

Similar tests can also be performed which indicate the comparability of these parameters between different groups of data.

Problems arise, however, in the interpretation of certain of the results of the tests. This is true particularly in the instance where an observed correlation coefficient is tested for a significant difference from unity (or 0.99). Should the observed value of r be found

significantly smaller than unity, it is concluded that the distribution of settlements is not consistent with that specified by central place theory. That is, the various sizes of places are not uniformly distributed. But, to label the observed spatial pattern "random," would be incorrect, since r can also be tested for a significant difference from zero. In a large sample it is not unlikely that the correlation coefficient is both significantly different from unity and zero, a situation which presents somewhat of a barrier to certain types of evaluation, although not to a comparison with central place theory.

Future research in the use of these and similar inferential models requires further definition and refinement of terms and techniques. Artificially generated locations and size distributions are necessary for this type of undertaking. Modern, large capacity computers should prove most useful in this type of research, since their large memories and high computing speeds bring such tasks within the reach of the geographer.

Finally, empirical research should be carried on using a variety of different types of point phenomena. Central place theory provides a hierarchical relationship in which the locations of settlements are related to their position in the array of functional activities and to their immediate surroundings. Other discussions of functional

relationships include interdependencies among settlements, but fail to provide definitive statements as to the spatial arrangement of the cities involved. Analysis of the spatial characteristics of functional nodes and their relevant consuming populations may be possible with the further development of the techniques proposed herein. Analyses of the distribution of completely independent phenomena, such as of centers of lunar craters of varying sizes, can also be useful in the definition of randomness in observable situations.

APPENDIX A

THE FRACTILE DIAGRAM

The entire framework of this research is based on that group of mathematical techniques known as "inferential statistics." This set of techniques is particularly useful to the researcher, since it allows him to make certain statements concerning his work which the more common descriptive statistics do not. Through the use of inferential methods, the researcher is capable of testing for significant differences among different groups of data, and is allowed to discuss these differences with certain degrees of confidence.

One of the more common requirements of inferential statistics is that the data be of a normal form. That is, most of the inferential techniques assume that the distribution of the data about the mean of the observations will conform to a rigid pattern -- that of the normal curve. The fractile diagram is one technique which can be used to test the degree of conformity between a collection of observed data and the form which is required for inferential analyses.

The principle under which the fractile diagram is formulated is quite simple. Since the normal distribution is symmetric about the mean, by definition one half of the observations are less than the mean and one half are greater. The mean of a normal distribution, therefore, is equal to the median, and is the fiftieth percentile in rank. All other values in a normal distribution are also defined as to their ranks. For example, that value which is one standard deviation greater than the mean has a percentile rank of 84.9, whereas 15.1 is the percentile of that value one standard deviation less than the mean. The fractile diagram can be used as a test of normality by comparing that value which is observed to have a particular percentile rank with the theoretically expected value. A partially symbolic discussion follows.

Assume the existence of a collection of observed data, X. The individual observations may or may not be arranged in any particular order, but they and the entire population may be characterized as follows:

X = an arbitrary collection of data;

N = the number of observations in X;

x = any single observation within X;

x_i = the i^{th} observation within X;

$\overline{X}$ = the arithmetic mean of X;

$S_X^{\,2}$ = the variance of X;

S_X = the standard deviation of X

$Z_{x_i} = (x_i - \overline{X}) / S_X$

In addition to this information, each x_i must be ranked according to its size, such that the largest observation has the greatest rank and the others are ranked in order of descending size. In this arrangement, then,

P_{x_i} = percentile rank of the i^{th} observation of X,[56]

$Q_{x_i} = 1 - P_{x_i}$.

In the case that several observations are equal, the percentiles associated with each are equal by definition. But, rather than assign each observation a rank of

$(l_i - 1/2) / N$,

---

[56]Because of operational requirements (neither P_{x_i} nor $1-P_{x_i}$ may equal 0), the percentile rank of the i^{th} observation is not defined as

l_i / N,

but rather as

$(l_i - 1/2) / N$,

in which

l_i = the number of observations which are less than or equal to the i^{th} observation.

This definition provides that the largest percentile rank, max P_{x_i}, is something less than unity, and that

$$P_{x_i} - 0 = 1 - (0.50 + (0.50 - P_{x_i})).$$

Thus, the distribution of P_X is symmetrical about 0.50 and does not approach the limiting values of 0 and 1 in anything other than simultaneous, balanced increments.

which is the high extreme of a whole range of values which are pre-empted by the presence of equal observations, each is assigned the mean of this set of percentile values, this mean percentile, $\overline{P}_{x_i}$, is defined as

$$(((l_i - 1/2) / N) + \sum_{j-1}^{n-1} (((l_i - 1/2) - j) / N)) / n,$$

in which

n = the number of observations equal to the i^{th} one.

The use of the fractile diagram also requires the definition of several terms associated with the theoretical normal distribution.

P = the percentiles associated with a normal distribution in which N observations are ranked:

P_j = the j^{th} percentile within P;

x_j = the j^{th} value in a $N(\overline{X}, S_x)$ distribution.

x_{p_j} = that x_j value associated with a particular percentile, P_j, such that

$Z_{p_j} = (x_{p_j} - \overline{X}) / S_x$, and

$$\varphi_{p_j} = \frac{1}{2} e^{\frac{-Z_{p_j}^2}{2}}$$

These values are used in testing a collection of observed data, X, for conformity to a normal distribution according to the method described by Hald.[57]

---

[57] Anders Hald, "The Normal Distribution," Statistical Theory with Engineering Applications (New York: John Wiley and Sons, Inc., 1952), pp. 119-43.

For any distribution of data where $P_{x_i} = P_j$, x_i may vary from x_{p_j} only randomly, and must fall within certain confidence limits. The distribution of x_i about x_{p_j} is assumed to be normal, and limits are set up to include a certain portion, say 90%, of the probable values about x_{p_j}. If x_i fails to fall within the specified acceptance region, it is judged to be significantly different from x_{p_i} at the 90% level of confidence. Should more than ten percent of the x_i's be found to be significantly different from their respective x_{p_j}'s, the entire distribution of observed data is rejected as being significantly different from normal.

The confidence limits about x_{p_j} are determined as follows, using 90% as a desired level:

Given x_i and its rank P_{x_i}, the variance of x_i about x_{p_j} ($P_j = P_{x_i}$) is defined as

$$S_{x_j}^{\ 2} = \frac{P_j \cdot Q_j}{N \cdot \varphi_{p_j}^{\ 2}} .$$

The standard deviation of course, is the square root of $S_{x_i}^{\ 2}$, or S_{x_j}. And, since 90 percent of all values of a normal distribution lie within 1.645σ of the mean, substituting x_{p_j} for the mean and S_{x_j} for σ, x_i must fall within the range $x_{p_j} \pm 1.645 \cdot S_{x_j}$ in order to be considered a normally distributed observation within the specified confidence limits.

The data used in this dissertation was tested for normality using a computer program based on the method described above.[58]

The fractile diagram technique is useful for another reason. After a collection of data has been demonstrated to conform to normality, that is, the x_i's vary only randomly from the x_{p_j}'s, where $P_{x_i} = P_j$, "equal size" categories may be established. These categories are of particular interest, because they do not destroy the continuous character of such a distribution. This quality is of great value in research the likes of this dissertation.

The equal size categories are not class intervals in the usual sense of the word, but are set up about each observed value, x_i. This is accomplished by substituting x_i for x_{p_j}, and considering all values of X which fall within the range of $x_i \pm 1.645 \cdot S_{x_j}$ as being equal to x_i. Thus, there are as many equal size categories as there are different values of x_i. None of these categories is mutually exclusive, and in the same way that the confidence

---

[58]Clifford E. Tiedemann, Program FRACTL (Los Angeles: Department of Geography, University of California at Los Angeles, 1965). This routine is also being prepared for submission to the SHARE library of computer programs, and will be available to subscribers within several months.

limits about the x_{p_j}'s expanded as P_j got larger or smaller so do the equal size limits about the x_i's. Equal size categories about each observed value were established using a computer program based on this substitution.[59]

The array of data being tested for normality need not consist of raw observations. Indeed, in many cases where 0 is a limit on the range of possible observed values, some transformation is necessary to render the data normal. This was the situation in the instance of the populations of Michigan settlements. The data transformation used was as follows:

$$x_i = \sqrt{\log_{10} \text{population}_i}.$$

This transformation produced a mean of 1.53205 and a standard deviation of 0.26990 (see Figure III-1). Of the 1062 places with populations larger than 100, approximately 91.8 percent fell within the 90 percent confidence limit. Since less than ten percent of the x_i's were found to be significantly different from their respective x_{p_j}'s, the distribution of transformed values were accepted as being normal.

---

[59]Clifford E. Tiedemann, Program EQUAL (Los Angeles: Department of Geography, University of California at Los Angeles, 1965). This routine is also being prepared for submission to the SHARE library of computer programs, and will be available to subscribers within several months.

In the case of the standardized distance, no transformation was found which proved to be successful. It was finally decided that the one which came closest (61.06 percent inside the ninety percent confidence limits) was to be adopted (see Figure III-2). This transformation had the form

$$x_i = \sqrt[3]{\text{standardized distance}_i}.$$

APPENDIX B

FORMULAE USED IN THE COMPUTATION OF DISTANCE BETWEEN URBAN PLACES

Because of the number of observations being considered in this research, approximations of road distances between each place and its order neighbors of equal or larger size are used. In order to do this, two situations must be considered in making the calculations. First, it is possible that the points might be connected by a fairly straight road passing through both of them. Second, since much of Michigan is served by section line roads, it was considered quite possible that two such places might be connected by a route running east-west and north-south and possessing a right angle intersection. Therefore, two different distances were calculated for each pair of points, one straight-line and one right-angle distance.

In addition, the fact that Michigan is made up of two peninsulas created another problem. In order to account for the fact that the only connection for land travel between the two is by way of the Mackinac Bridge, straight line distances between points on opposite peninsulas were

computed in two segments -- one from place i to the bridge, and the second from the bridge to place j. (It is shown that only two cities had their nearest neighbor on the opposite peninsula: Saint Ignace was the nearest neighbor for Mackinac City, and Cheboygan was Saint Ignace's nearest neighbor.)

The approximated distances were calculated for both major categories, straight-line and right-angle, and were ranked in order of ascending size. Then, the smaller of these order neighbor distances for each settlement were visually compared with the map situation to see which best depicted the road situation. The one was considered to be most representative was selected as the value to be used in the statistical analyses of the size-distance relationships of Michigan settlements.

The formulae used were standard half-angle trigonometric functions. In the instance of straight distances between places i and j on the same peninsula, the formula was as follows:

$$\sin(\text{theta} / 2) = ((\cos(\text{lat}_i) \cdot \cos(\text{lat}_j) \cdot \sin^2(|\text{lon}_i - \text{lon}_j| / 2)) + \sin^2(|\text{lat}_i - \text{lat}_j| / 2))^{1/2},$$

in which

theta = the distance between places i and j in radians,

lat_i = the latitude of place i,

lat_j = the latitude of place j,

lon_i = the longitude of place i, and

lon_j = the longitude of place j.

In order to find the distance between i and j in terms of miles, it is necessary to multiply theta by the radius of the earth at an appropriate latitude.

In the case of right-angle distance between places on the same peninsula, a modification of the previous formulae were used,

$$\sin(dlon / 2) = (\cos^2((lat_i + lat_j) / 2) \cdot \sin^2(|lon_i - lon_j| / 2))^{1/2} \text{ and}$$

$$\sin(dlat / 2) = \sin(|lat_i - lat_j| / 2) ,$$

in which

dlon = the relevant east-west distance in radians, and

dlat = the relevant north-south distance in radians.

Because of the convergence of meridians toward the poles, the east-west distance was calculated along a parallel midway between those of the two cities.

The straight-line distance by way of Macknica Bridge was calculated by using the straight-line formula twice. They had the following appearance:

$$\sin(theta_i / 2) = ((\cos(lat_i) \cdot \cos(lat_b) \cdot$$

$\sin^2(|lon_i - lon_b| / 2)) + \sin^2(|lat_i - lat_b| / 2))^{1/2}$

and

$$\sin(theta_j / 2) = ((\cos(lat_b) \cdot \cos(lat_j) \cdot \sin^2(|lon_b - lon_j| / 2)) + \sin^2(|lat_b - lat_j| / 2))^{1/2},$$

in which

$theta_i$ = the straight-line distance between place i and the bridge in radians,

$theta_j$ = the straight-line distance between the bridge and place j in radians,

lat_b = the latitude of the Mackinac Bridge, and

lon_b = the longitude of the Mackinac Bridge.

The right-angle distance between places on opposite peninsulas involved four formulae, each a modification of the basic straight-line function. They appeared as below:

$$\sin(dlon_i / 2) = (\cos^2((lat_i + lat_b) / 2) \cdot \sin^2(|lon_i - lon_b| / 2))^{1/2},$$

$$\sin(dlat_i / 2) = \sin(|lat_i - lat_b| / 2),$$

$$\sin(dlon_j / 2) = (\cos^2((lat_b + lat_j) / 2) \cdot \sin^2(|lon_b - lon_j| / 2))^{1/2}, \text{ and}$$

$$\sin(dlat_j / 2) = \sin(|lat_b - lat_j| / 2),$$

in which

$dlon_i$ = the relevant east-west distance between place i and the bridge,

$dlat_i$ = the relevant north-south distance between place i and the bridge,

$dlon_j$ = the relevant east-west distance between the bridge and place j,

$dlat_j$ = the relevant north-south distance between the bridge and place j.

These computed distances are then standardized by multiplying them by the square root of the density associated with place i. This i^{th} density is calculated by dividing the number of settlements equal to or larger than the i^{th} city by the area of the state.

In order to compare the observed size-distance relationship with that which might be expected according to central place theory, it is necessary to calculate expected standardized distances. This is done using the following formula:

$$E_{i_j} = 1.075 \sqrt{A / n_i} ,$$

in which

E_{i_j} = the expected real distance between place i and its nearest neighbor, place j,

A = the area of the study region (58,216 in the case of Michigan), and

n_i = the number of places equal to or larger than place i.

This expected distance is standardized in the same manner as are the observed distances. That is,

$$\text{standardized } E_{i_j} = (1.075 \sqrt{A / n_i}) \cdot (n_i / A)^{1/2},$$

or

$$\text{standardized } E_{i_j} = 1.075.$$

BIBLIOGRAPHY

Anderson, Theodore R. "Potential Models and Spatial Distribution of Population," Papers and Proceedings, Regional Science Association. II (1956), pp. 175-82.

Arkin, Herbert and Raymond R. Colgon. Tables for Statisticians: College Outline Series. New York: Barnes and Noble, Inc., 1963.

Artle, Roland. "On Some Methods and Problems In the Study of Metropolitan Economics," Papers and Proceedings, Regional Science Association. VIII (1962), pp. 71-88.

Bachi, Roberto. "Standard Distance Measures and Related Methods for Spatial Analysis," Papers and Proceedings, Regional Science Association. X (1963), pp. 83-132.

Beckman, Martin J. "City Hierarchies and the Distribution of City Size," Economic Development and Cultural Change. VI, No. 3 (April, 1958), pp. 243-8.

Berry, Brian J. L. "The Impact of Expanding Metropolitan Communities Upon the Central Place Hierarchy," Annals, Association of American Geographers. L, No. 2 (June, 1960), pp. 112-6.

________ and H. Gardiner Barnum. "Aggregate Relations and Elemental Components of Central Place Systems," Journal of Regional Science. IV, No. 1 (Summer, 1962), pp. 35-68.

________, ________ and Robert J. Tennent. "Retail Location and Consumer Behavior," Papers and Proceedings, Regional Science Association. IX (1962), pp. 65-105.

________ and William L. Garrison. "Recent Developments of Central Place Theory," Papers and Proceedings, Regional Science Association. IV (1958), pp. 107-20.

________ and ________. "A Note on Central Place Theory and the Range of a Good," Economic Geography. XXXIV, No. 4 (October, 1958), pp. 304-11.

________ and ________. "The Functional Bases of the Central Place Hierarchy," Economic Geography. XXXIV, No. 2 (April, 1958), pp. 145-54.

________ and ________. "Alternate Explanations of Urban Rank-Size Relationships," Annals, Association of American Geographers. XLVIII, No. 1 (March, 1958), pp. 83-90.

________ and Allen Pred. Central Place Studies: A Bibliography of Theory and Applications. Philadelphia: Regional Science Research Institute. 1961.

Blome, Donald A. "An Analysis of the Changing Spatial Relationships of Iowa Towns, 1900-1960," (unpublished Ph.D. dissertation) Iowa City: State University of Iowa, 1963.

Bogue, Donald J. and Calvin L. Beal. Economic Areas of the United States. New York: The Free Press of Glencoe, Inc., 1961.

Brush, John E. "The Hierarchy of Central Places in Southwestern Wisconsin," Geographical Review. XLIII, No. 3 (July, 1953), pp. 380-402.

Bunge, William. Theoretical Geography: Lund Studies in Geography. Series C. General and Mathematical Geography, No. 1. Lund: Department of Geography, The Royal University of Lund, 1962.

Bureau of the Census. "Table 8., Populations of All Incorporated Places and Unincorporated Places of 1000 or More, 1960," United States Census of Population, 1960: Michigan: Number of Inhabitants PC(1)24A. Washington: United States Department of Commerce, 1961, pp. 23-5.

Carol, Hans. "Hierarchy of Central Functions Within the City," Annals, Association of American Geographers. L, No. 4 (December, 1960), pp. 419-38.

Christaller, Walter. Central Places in Southern Germany. (Carlisle W. Baskin, trans.) Englewood Cliffs: Prentice Hall, Inc., 1966.

Clark, Philip J. and Francis C. Evans, "Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations," Ecology. XXXV, No. 4 (October, 1954), pp. 445-53.

Dacey, Michael F. "A Note on Some Number Properties of a Hexagonal Hierarchical Plane Lattice," Journal of Regional Science. V, No. 2 (Winter, 1964), pp. 63-9.

________. "Two Dimensional Point Patterns: A Review and Interpretation," Papers, Regional Science Association. XIII (1964), pp. 41-58.

________. "Imperfections in the Uniform Plane," Discussion Paper No. 4 of the Michigan Interuniversity Community of Mathematical Geographers. Ann Arbor: Department of Geography, University of Michigan for the Michigan Interuniversity Community of Mathematical Geographers, June, 1964. (mimeo.)

________. "Analysis of Central Place and Point Patterns by a Nearest Neighbor Method," Proceedings of the I.G.U. Symposium on Urban Geography, Lund, 1960: Lund Studies in Geography, Series, Human Geography, No. 24. (Knut Norborg, ed.). Lund: Department of Geography, Royal University of Lund, 1962, pp. 55-76.

Dixon, Wilfrid J. and Frank J. Massey, Jr. Introduction to Statistical Analysis. New York: McGraw-Hill Book Co., Inc., 1957.

Geological Survey. "Michigan," National Topographic Map Series, 1:250,000. Washington: United States Department of the Interior, 1954-64.

Getis, Arthur. "Temporal Land-Use Pattern Analysis With the Use of Nearest Neighbor and Quadral Methods," Annals, Association of American Geographers. LIV, No. 3 (September, 1964), pp. 391-99.

Gibbs, Jack P. "The Evolution of Population Concentration," Economic Geography. XXXIX, No. 2 (April, 1963), pp. 119-29.

Hald, Anders. "The Normal Distribution," Statistical Theory with Engineering Applications: Wiley Publications in Statistics. (G. Seidelin, trans.) New York: John Wiley and Sons, Inc., 1952, pp. 119-43.

Hoover, Edgar M. "The Concept of a System of Cities: A Comment on Rutledge Vining's Paper," Economic Development and Cultural Change. III, No. 2 (January, 1955), pp. 196-8.

________. *The Location of Economic Activity*. New York: McGraw-Hill Book Co., Inc., 1948.

Isard, Walter. *Location and Space Economy*. New York: John Wiley and Sons, Inc., 1958.

King, Leslie J. "A Multivariate Analysis of the Spacing of Urban Settlements in the United States," *Annals*, Association of American Geographers. LI, No. 2 (June, 1961), pp. 222-33.

Lösch, August. *The Economics of Location*. (W. H. Woglom and W. F. Stolper, trans.) New Haven: Yale University Press, 1954.

________. "The Nature of Economic Regions," *Southern Economic Journal*. V, No. 1 (July, 1938), pp. 71-8.

"Michigan: Index to Cities, Towns, Counties, Transportation Lines, Airports, Banks, and Post Offices," *Commercial Atlas and Marketing Guide: 1962*. Chicago: Rand McNally and Co., 1962, pp. 233-7.

Michigan: Official Highway Map: 1961. Lansing: Michigan State Highway Department, 1961.

Morrill, Richard L. "The Development of Spatial Distribution of Towns in Sweden: An Historical-Predictive Approach," *Annals*, Association of American Geographers. LIII, No. 1 (March, 1963), pp. 1-14.

________. "Simulation of Central Place Patterns Over Time." *Proceedings of the I.G.U. Symposium on Urban Geography, Lund, 1960: Lund Studies in Geography, Series B. Human Geography, No. 24*. (Knut Norborg, ed.), Lund: Department of Geography, Royal University of Lund, 1962, pp. 109-20.

Nelson, Howard J. "Some Characteristics of the Population of Cities in Similar Service Classifications," *Economic Geography*. XXXIII, No. 1 (January, 1957), pp. 95-108.

________. "A Service Classification of American Cities," *Economic Geography*. XXXI, No. 3 (July, 1955), pp. 189-210.

Philbrick, Allen K. "Principles of Areal Functional Organization in Regional Human Geography," Economic Geography. XXXIII, No. 4 (October, 1957), pp. 299-336.

________. "Areal Functional Organization in Regional Geography," Papers and Proceedings, Regional Science Association. III (1957), pp. 87-98.

Schnor, Leo F. "The Functions of Metropolitan Suburbs," Papers and Proceedings, Regional Science Association. V (1959), pp. 272-9.

________. "Satellites and Suburbs," Social Forces. XXXVI, No. 4 (December, 1957), pp. 121-7.

Stewart, Charles T., Jr. "The Size and Spacing of Cities," Geographical Review. XLVIII, No. 2 (April, 1958), pp. 222-45.

Thomas, Edwin N. "The Stability of Distance-Population-Size Relationships for Iowa Towns from 1900-1950," Proceedings of the I.G.U. Symposium in Urban Geography, Lund, 1960: Lund Studies in Geography. Series B. Human Geography, No. 24. (Knut Norborg, ed.), Lund: Department of Geography, Royal University of Lund, 1962, pp. 3-12.

________. "Toward an Expanded Central Place Model," Geographical Review. LI, No. 3 (September, 1961), pp. 400-11.

Ullman, Edward L. "A Theory of Location for Cities," American Journal of Sociology. XLVI, No. 6 (May, 1941), pp. 853-64.

Vining, Rutledge. "A Description of Certain Aspects of an Economic System," Economic Development and Cultural Change. III, No. 2 (January, 1955), pp. 147-95.

ERRATA

Page 30, Figure II-1. "discreet" should read "discrete"

Page 58, Figure III-4. "SHORLINE" should read "SHORELINE"

POPULATIONS are transformed by the factor $\sqrt{\text{Log}_{10}\text{ population}}$

STANDARDIZED DISTANCES are transformed by the factor

$$\sqrt[3]{\text{standardized distance}}$$

Page 71, Figure III-9. POPULATIONS are transformed by the factor $\sqrt{\text{Log}_{10}\text{ population}}$

STANDARDIZED DISTANCES are transformed by the factor

$$\sqrt[3]{\text{standardized distance}}$$

109