A TECHNIQUE FOR PREDICTING THE URBAN STRUCTURE OF A REGION

by

WILLIAM MICHAEL WARWICK

A RESEARCH PAPER

submitted to

THE DEPARTMENT OF GEOGRAPHY

in partial fullment of the
requirements for the
degree of

MASTER OF SCIENCE

June 1976

Directed by
Dr. R. M. Northam

ABSTRACT

The two classic theories of the distribution of cities in an urban system are the rank size rule and central place theory. The rank size relationships, as incorporated into G. K. Zipf's Principle of Least Effort, is seen as the result of a balancing of the opposing forces of urban location. Central place theory, as developed by W. Christaller, presents an alternate view of cities as ordered groups of cities with the cities of each group sharing unique characteristics. Because the rank size distribution is a continuous one and central place theory rests on the assumption of discretely distributed clusters of cities of similar population, the two theories are considered to be incompatable. The present paper summarizes the two theories as developed by G. K. Zipf and W. Christaller, discusses the conflicts between the two theories, and develops a semilogarithmic graphing technique that provides a methodology for resolving the conflicts between them. When rank size distributed data are plotted on semilogarithmic graph paper, and the resulting curve approximated by a series of straight lines, groups of cities characteristic of the central place distribution are defined. Linear regressions on these data reveal consistant parameters characteristic of central place distributions. These parameters are dimensionless and have been found to be useful in the comparative analysis of urbanized regions. They may also be useful in the analysis of urban structure and correlated phenomena.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The two classic theories of the distribution of cities in an urban system are the rank size rule and central place theory. The linear rank size relationship is revealed when cities are ranked by relative population and plotted on logarithmic graph paper. This relationship, as explained by G. K. Zipf as the Principle of Least Effort, is the result of the balancing of opposing forces of urban location.[1] Central place theory, as developed by W. Christaller describes the distribution of cities in a region in terms of a discrete step like hierarchical distribution of cities grouped into orders.[2] Cities within these orders theoretically share characteristics of population, function, territory, and trade areas. Because the rank size rule generalizes from a continuous geometric distribution of city populations and central place theory is based on a discrete geometric frequency distribution of urban orders, these two theories and their conclusions have been considered to be incompatible. It is the objective of this paper to summarize these two theories as they were developed by G. K. Zipf and W. Christaller, to comment on the theoretical differences that have been noted, and to describe a methodology for resolving these apparent conflicts. The new model of urban structure within a region derived from this analysis provides a new perspective of urban systems and a number of dimensionless parameters that may be useful in the comparison of urbanized regions and correlated phenomena.

TWO THEORIES OF URBAN SYSTEMS

Numerous economic, sociologic, and geographic studies have attempted to quantify the observed differences that exist between cities. These attempts have led to the empirical generalizations of the urban rank size rule and of central place theory. The Principle of Least Effort as developed by G. Zipf uses a special application of the rank size rule to characterize the optimization of economic forces in an urban system using the empirical regularities that can be observed in the distribution of cities in the system. Central place theory, as developed by W. Christaller, attempts to explain the population, number, and location of cities using indicies of communication, transportation, and exchange between cities. This theory concentrates on understanding the functional character of individual places and their subsequent distribution. Places of similar character are treated as members of a distinct order and are thereby related to all other orders in a hierarchical fashion. Since the model developed in this paper includes elements of both of these theories, a more detailed summary is necessary.

Zipf's "Principle of Least Effort"

In any large area, it may be observed that there are few large cities, more numerous medium sized cities, and still more numerous smaller sized places. When these places are ranked in order according to population and the resulting points plotted on doubly logarithmic graph paper, population on the Y axis and rank of the place on the X axis, a nearly linear relationship emerges (Appendix A is representative of this technique). Generalized as the rank size rule, this relationship varies from region to region and holds best for cities above an arbitrary minimum population. Although Zipf was not the first to observe rank size regularities in urban size distributions, he was the first to systematically study the phenomenon and develop a comprehensive theory to explain the relationship.[3]

Zipf's theory of the Principle of Least Effort as he applied it to urban phenomena incorporated three fundamental concepts: the limited domain of a good, the forces of unification, and the forces of diversification. He also recognized the differential ability of goods to bear the cost of manufacture, marketing, and transportation. Since he realized that only certain goods could be sold and distributed throughout a large region from only one point, he developed a theory of economic and locational dynamics to explain how all goods might be manufactured and distributed in a number of different patterns.

According to Zipf, urban rank size regularities are the result of the interaction of the competing forces of diversification and unification. The forces of diversification are the totality of forces, such as point resources and limited concentrations of labor, working to tie production to many scattered locations. Early manufacturing economies are dominated by these forces and result in a range of similar goods being produced and consumed at numerous relatively self-contained places throughout the territory. The strength of this force is generally indicated by the number, distribution, and population of small places. It is represented graphically by a long tail on the rank axis of the distribution. In similar study areas, the total number of places could act as a crude index of this force. The force of unification represents the totality of forces such as economies of scale, and transport economies, acting to concentrate population, production, and consumption in fewer larger cities. It is represented graphically by a long tail on the population axis. Among similar study regions, the population of the largest city could act as a rough index of the force of unification.

As part of his research, Zipf ranked the one hundred largest metropolitan districts in the United States in 1940 in descending order of population. These data were plotted on doubly logarithmic graph paper, population on the Y axis and rank on the X axis. This distribution formed a slightly concave curve that was easily approximated by a straight line with a slope near unity. According to this theory, a slope of one is the natural result of a balancing of the forces of diversification and the forces of unification. He generalized this relationship into a family of equations based on the function for geometric series, $Y = aX^b$, or, in his notation, $k = r \cdot p^q$. In his equation, the two parameters k and q represent the population of the largest city and the ratio of the forces of diversification divided by the forces of unification respectively. The two variables r and p represent the values of city rank and population, X and Y, respectively. This type of exponential equation is known as the geometric or allometric equation and this type of continuous distribution is central to the rank size rule.[4]

Christaller's Central Place Theory

G. K. Zipf's Principle of Least Effort generalized from the rank size regularities of a regional urban system to the dynamics and character of that region. W. Christaller builds instead on individual places and the relationships between places in constructing his theory of regional structure known as central place theory.[5] In this theory, the primary function of a city is to act as a center for the distribution of goods consumed but not produced locally. Each of these distribution centers, or central places, has associated with it a tributary area with consumers who find that particular center to be the most convenient market for the goods and services available. As a result of this relationship, and of the competition of other centers serving similar functions nearby, the central place is located at or near the center of its tributary area at the point of minimum aggregate travel.

Christaller observed a frequency distribution of places of various populations similar to Zipf's rank size distribution. Additionally, he noticed that the numerous small places carried stocks of similar widely demanded goods and services. Goods and services that were infrequently demanded by the local population were rare. However, at places of the next larger order, goods and services not readily available at smaller places were more generally available in addition to the same spectrum of goods and service that were available at the lower order centers. Christaller observed that goods and services of different demand, profitability, and domain were differentally associated with central places of a corresponding order. Each order formed a discrete, recognizable class of cities of similar population, market area, and consumer accessibility, or centrality. These various orders of central places were, in turn, related to each other in a hierarchical fashion.

Christaller identified these different orders of central places in his research in Southern Germany. In order of increasing centrality, they are the M, A, K, B, G, P, L, RT, and R orders. Although he identified the RT and R orders, he described them as unique L order central places that have taken on the additional functions of national center or world city and are not likely to occur in most regional urban systems.[6] He considered the developed L system (the complete hierarchy of places of orders M through P, ending in a single city of the L order), the normal system of regional urban organization. Because of its role at the apex of the hierarchy, and its dominance as the largest single city, the L order place acts as the main channel for interregional circulation and exchange. Each of these regional centers competes with the others for its region's share of interregional trade goods and acts as the broker for its region's goods and services. As a result, a nation can be described as a collection of relatively self-contained regions articulated by L order regional capitals.

Christaller also observed within each region a characteristic pattern of organization and location for the places of that region. If the region is assumed to be on a uniform plain with equal access from all directions, the market areas of adjacent central places of the same order form six-sided tributary territories. This pattern results from the process of economic competition because each place is more convenient for local consumers than any other. Given Christaller's original assumption about the uniformity of population for all places of a given order, the area and length of each side of these hexagonal market areas are also uniform and are characteristic of the places of that order. If we can imagine a dense network of uniform hexagonal territories on a uniform plain with low order places at the center of each hexagonal trade area, we can reconstruct the pattern of organization that Christaller observed. If six adjacent places of this same order are seen as forming a larger hexagon with a place on each point they would then describe the hexagonal trade area of a place of at least the next higher order. If six places of this higher order are grouped into a still larger hexagon with a place on each point, they would delineate yet another trade area for a still higher order place, and so on (Fig. 1).

Within his generalized hexagonal scheme of central place locations on a uniform plain, Christaller identified three distributional principles. Each of these principles results in a characteristic pattern for nesting of adjacent places and for transportation routes. According to his marketing principle, adjacent places are nested in triangular clusters of three, regardless of order. These clusters are in turn nested in clusters of three and so on (Fig. 2, upper diagram). If political administrative principles dominate regionally, the lower order places are nested in clusters of seven with a higher order administrative place at the center of this six-sided cluster. Each of these clusters is in turn nested with seven adjacent clusters with a still higher order place at the center and so on (Fig. 2, middle diagram). If transport economies dominate regionally, nesting along linear transportation routes occurs in clusters of four in a diamond pattern (Fig. 2, lower diagram).

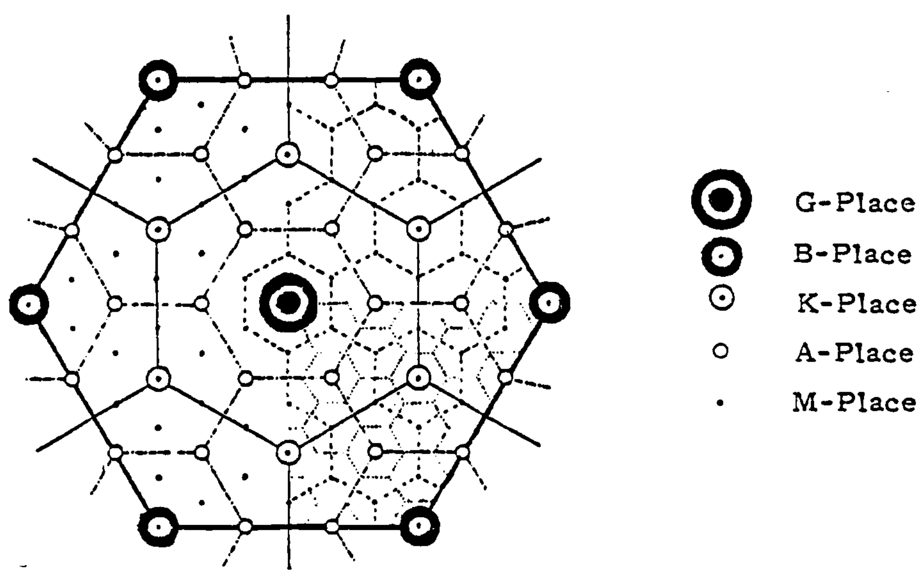


Fig. 1. Christaller's hexagonal scheme for a system of central cities.

Source: Vining, op. cit, footnote 10.

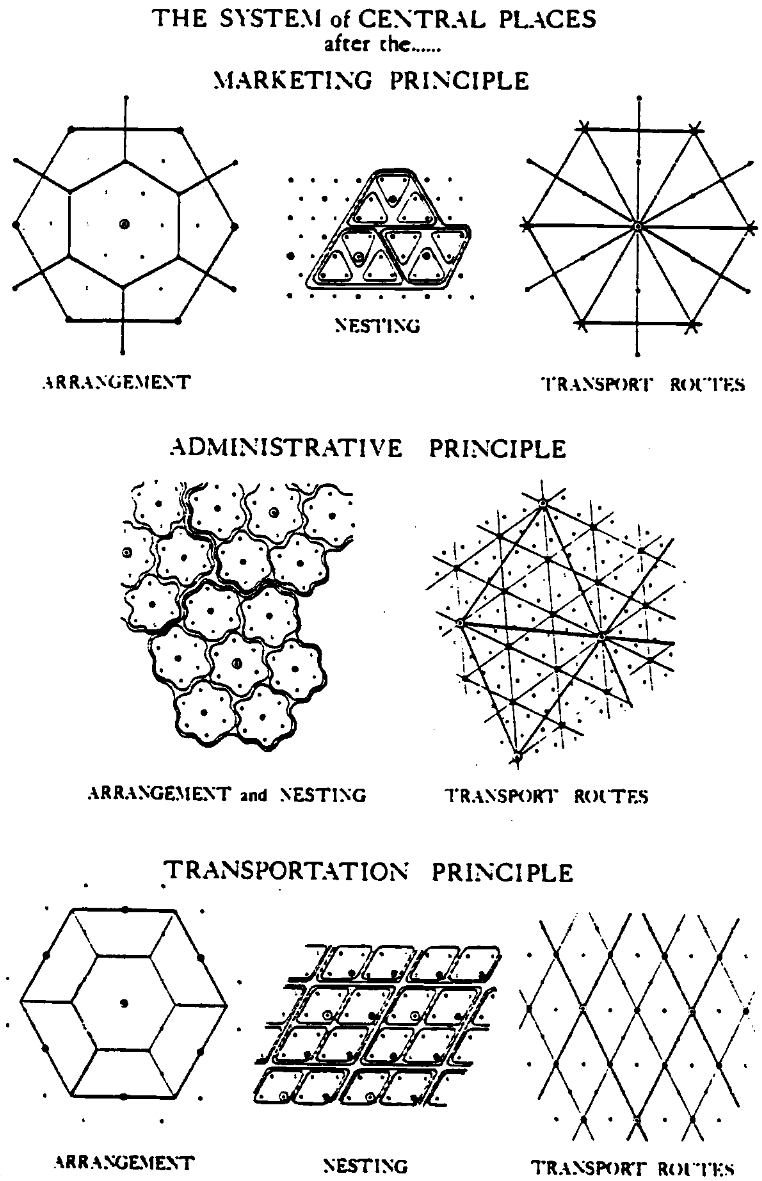


Fig. 2. Christaller's organizational patterns for the system of central places.

Source: Berry and Pred, op. cit, footnote 2.

In the marketing principle of regional spatial organization, the base for clustering is three; in the administrative principle it is seven, and in the transportation principle it is four. The following generalizations hold for all three cases, although only the K=3, or marketing principle, will be illustrated. In the K=3 system, the largest single city (the L order place) has a population of K. According to Christaller, the next lower order of places, the P order places, would consist of 3^1, or three places each with a population of $K/3^1$, or one third that of the largest city. Similarly, the next lower order of places would be 3^2 or 9, in number, and each place would have a population one ninth that of the largest city. This relationship extends similarly throughout the hierarchy. Thus, the total population of all places of a given order is equal to that of the largest city, and the population of each place of an order is a constant geometric proportion of the population of the largest city. This relationship can be generalized as an exponential equation of the form, $Y=aX^b$, where Y is the number of places of a given order b, a is a constant, (in this case a=1), X is the base of the system of organization, (3, 4, or 7), and b is the order of places you are trying to describe. From this discussion, it can be seen that in Christaller's theory, the number or frequency of places in an order increases exponentially as the order increases from L to M. It is important to remember that this exponential increase is based on the limited, discrete steps of the central place order hierarchy and as a result it is a limited discrete step-like increase.

Conflict between the Two Theories

Despite the popularity and empirical evidence in support of both of these theories, conflicts between them have been noted.[7] The central disagreement stems from the fact that the distribution of most urban phenomena, such as population, size of trade area, gross income, and so on, appear to follow a continuous rather than a discrete distribution as Christaller theorized in central place theory. Although Christaller's theory of central place clustering was primarily concerned with a discrete step-like clustering of central places based on the purchase of characteristic central goods, he assumed that correlated phenomena such as population and size of trade area would also follow this discrete distribution. Since continuity is apparent in the distribution of certain urban data, it is difficult to justify the differentiation and ordering of these same data. As a result, some of the schemes for classifying places into the various central place orders have been criticized as arbitrary and self-fulfilling. However, to the researcher familiar with both central place theory and his data, the deviations in apparently continuous distributions may reveal significant characteristics of the central place system.

Central place researchers are generally not provided with obvious indicators of the central place system in a region. Thus, they are forced to look for a reflection of the central place system in a variety of urban phenomena. As a result, some lack of clarity is to be expected. However, once a central place classification has been completed, parameters can be found to justify the classification and reveal the characteristic discrete ordering. Although this lack of obvious and consistant central place parameters makes the documentation and validation of central place systems difficult, the theory remains popular because of its intuitive appeal, internal consistancy, and descriptive power. However, despite the consistancy of the rank size distribution and its accurate description by an allometric formula, besides indicating some general kind of group membership it is relatively meaningless. This lack of significant theoretical meaning and descriptive power has been noted by B. Berry in his work with the rank size rule as well as by other researchers working with similar formulations.[8] In the following discussion, a method is presented to transform data that is distributed in rank size fashion into central place orders consistant with central place theory.

A TECHNIQUE FOR PREDICTING THE CENTRAL PLACE STRUCTURE OF A REGION

The models developed by Zipf and Christaller in support of their theories vary in data requirements and ease of application to field situations. Rank size models simply require reliable urban population data. These data are available, generally with some indication of their reliability, from national and international census organizations. Date requirements for central place models are more complex. The measures employed by Christaller included the population, size of tributary area, types of central functions provided, the number of central services available, and the distances between places for each place of each order.[9] These measures and generalized indices employed by others such as frequency of medical visits, telephone connections, airline connections, traffic volume, and so on, are not readily available and are virtually non-existent for purposes of historical research. Although urban population provides only a crude index of the importance and centrality of a place, it is an element of central place systems that is consistently measured and should reflect the central character of each place as well as the character of the central place system as a whole. A reliable technique utilizing population information would have great utility. The model developed in this paper utilizes rank size relationships and linear regression techniques to extract central place information from census population data in the construction of a central place model of a study area.

Sampling and Regional Analysis

Rather than proceeding directly to the model development section, let us first look at Figure 1. We can conduct an experiment that illustrates the importance of study area boundaries in central place and rank size surveys. If we use the two top and bottom sets of B-place points to delimit a rectangular study area, the count of places of each order represented will prove this point. Our count reveals 1, 4, 6, 14, and 42, G, B, K, A, and M order places, respectively. If we assumed the study area was a normal central place system, which it was, we would expect to find a distribution of 1, 3, 9, 27, and 81, G, B, K, A, and M order places in a K=3 system. Even if we knew we only had part of a system, we would be unable to revise our expectations because we would not know what part of the system we had sampled. This demonstration illustrates two points about central place studies: first, if a completely developed central place system is not included in the study area, the results will not accurately reflect the central place system and second, the larger an area and the greater proportion of the central place hierarchy you include in your study area, the less likely you will be to distort your sample of the central place system as it is developed. Since Christaller's central place hierarchy generally ends with an L place regional capital, complete regional study areas are most amenable to central place study. These same warnings apply to rank size surveys, although less emphatically. Because of these factors, the study area within which the following model was developed consisted of a complete regional system of cities.

Urban Structure in the Portland, Oregon Region: 1940-1970

Components of the Study

Since the study area had to contain a regional capital and its tributary region, a search of the regional science literature was undertaken to locate a previously defined area comprising such a system. In his 1971 central place research, R. Preston identified three such regions in the Pacific Northwest. Preston described complete regional hierarchies and tributary territories for each regional system of cities in his study area, including those of the regional capitals of Seattle, Washington and Portland, Oregon. Since the regions he described seemed appropriate for my analysis and inasmuch as his research could be useful in the verification of my research, I selected the region he associated with the Portland central place hierarchy as my study area.[10] This area was composed of the whole of Oregon State, (excluding Malheur County), the six border counties of southwestern Washington State, and a small border portion of northern California State south of Klamath Falls, Oregon.

Since the few places included in the California portion would not significantly contribute to the study, they were excluded; otherwise, my study area was identical to that described by Preston (Appendix B4). A similar study area could have been constructed using a population potential map or through use of a population gravity model.[11]

The technique discussed in the remainder of this paper was developed to indicate regional urban structure by which is meant: urban population distribution, hierarchy, orders of central places, and members of those orders. Since the primary data source was the decennial census for the years 1940, 1950, 1960, and 1970, and since this census does not accurately survey places of less than 1,000 inhabitants, this study was limited to analysis of places in excess of 1,000 inhabitants.[12] Generalizations however, extended to places below that figure.

Under the assumptions of Zipf's rank size theory, naturally bounded urban areas were required, coincidentally, these areas also provided his best theoretical fits. As a consequence, the populations of metropolitan districts or urbanized areas were used in my research whenever they were available. Frequently such aggregate data were not available and agglomerated urban area populations were estimated by combining the population of adjacent places and census divisions of significent population and density with the population of significantly large places.

Rank Size Regularities

The first step in the analysis of the study region was to plot the rank size distribution of all places over 1,000 inhabitants. According to Zipf's theory, this would give some indication of the homogeneity of the region and of internal balance among its various dynamic components. If the natural logarithm of the rank and the population are used, these data can be analyzed with linear regression techniques testing for conformance to the equation for a straight line: $Y=a+bx$, where a and b are constants, (a representing the place in the Y axis where the line intersects, and b indicating the slope of the line), and X and Y are variables. The approximation to a straight line is indicated by the coefficient of determination, r^2, which is calculated with another equation. The closer r^2 is to one, the better the approximation to a straight line.

Plots of the four data sets, one for each deciennial census, on double logarithmic paper revealed the characteristic linear distribution described by Zipf (Appendix A is representative of all four data sets). Linear regressions of the 100 largest places verified the observed linearity of these data (Appendix C). According to Zipf's theory, the slope parameter, b, would approximate one if the regions internal economic forces were balanced. In all four models, the slope parameters are greater than the expected values. Nevertheless, the regression models are all similar, indicating some sort of regional homogeneity and stability throughout the thirty year period under consideration. Since linear regression techniques are sensitive to wide variations in data and since the location of Portland on the doubly logarithmic graphs was so far out of line, another set of regressions were performed with Portland excluded. These regressions produced slightly lower slopes although the slopes all remained above unity (Appendix D). The size of Portland as predicted by both sets of regressions was substantially smaller than the observed population for Portland (Appendix E). We can make two generalizations from these regressions: first, because of the linearity of these data represented by the large r^2 factor and the consistency of the b parameter, this appears to be homogeneous region with some form of internal balance. Second, because the equations consistently predict a significantly smaller population for Portland, Portland must be excessively large for a region of this size. This large population would indicate that either Portland is a primate city and that the tributary region is underdeveloped (which should have been reflected in the second series of regressions), or that Portland's size is a result of central place functions that it provides to territories outside its tributary region.

Primacy in the Region

It has been observed in some underdeveloped nations that their urban network is dominated by one very large city, frequently ten times larger than the next largest place. Generally, a lack of intermediate size cities is evident, as is a lack of transportation and communication between intermediate size cities. This pattern has been associated with former colonies and with colonial powers as well. In the former, it is an indicator of underdevelopment and dependence while in the latter it is a remnant of former global strength. The excessively large cities which characterize this distorted urban network are called primate cities. B. Berry has developed a variation of the rank size plot which indicates primacy in an urban network. Since we have reason to suspect a primate distribution in the study area, we can utilize his simple technique.[13]

Berry's test for primacy uses lognormal probability graph paper. That is, paper with a probability scale on one side, the Y axis, and a logarithmic scale on the other, the X axis. The cumulative percentage of places above an arbitrary size are plotted on the probability axis and the population of the places is plotted on the logarithmic axis. In the article describing this technique, Berry aggregated his data and divided it into six size classes, ranging from 20,000 to over one million population. Berry's size classes varied significantly and inconsistently, so the technique was modified to utilize a continuous distribution of data from the study area. Analysis employing this modified technique produced a linear plot which indicated no tendency to primacy in the region (Appendix F is representative of all four data sets). If Portland were a primate city, it would not lie on the trend line formed by the other cities. In view of this test, Portland's apparently excessive population must be the result of its providing central functions to areas outside its tributary area via its port facilities. It is likely that Portland functions as a central place for Pacific Rim nations as well as for the people of Alaska, Hawaii, and parts of coastal California and Washington States. Extra regional central functions of this type would be reflected in Portland's population but would not be reflected in either the population or the distribution of intermediate and lower order places. The nearly constant slopes observed on the cumulative frequency graphs also indicate that the study region is a valid, coherent region which maintains a consistent internal balance through time. However, they leave unanswered the question of structure within the region.

Semi-logarithmic Rank Size Characteristics

If there are hierarchical orders of places in a region, as central place theory states, the forces operating on each order should be reflected differently according to the population of places. Except in the case of Portland, analysis of doubly logarithmic graphs does not readily reveal gaps, irregularities, or slope variations which could indicate such intraregional structure. However, if the same information is plotted on semi-logarithmic paper, with the logarithm of population on the Y axis and the rank of each place on an arithmetic scale on the X axis, a different interpretation of these data is possible.

A straight line on semi-logarithmic graph paper describes an exponential relationship described by the general formula, $Y=ae^{bX}$. This equation is called the exponential equation because it uses a variable in the exponent. It should not be confused with other exponential equations such as the geometric or allometric equation employed by Zipf. When plotted on semi-logarithmic graph paper, these data from the study area form an irregular curved line (Fig. 3 is representative of all four data sets). Distributed in this fashion, these data reveal gaps and variations in slope. If these curves are approximated by a series of straight lines, excluding Portland, they can be broken into multiple straight line segments (Figure 3, fitted lines indicate line segments). Linear regression analysis and least squares curve fitting techniques were used to confirm, strengthen, and define these apparent linear relationships.

Linear regression analysis of the points contained within each of the observed line segments was undertaken to clarify the apparent central place ordering indicated by the linear associations. Numerous regressions were performed by varying sample sizes to include or exclude points on

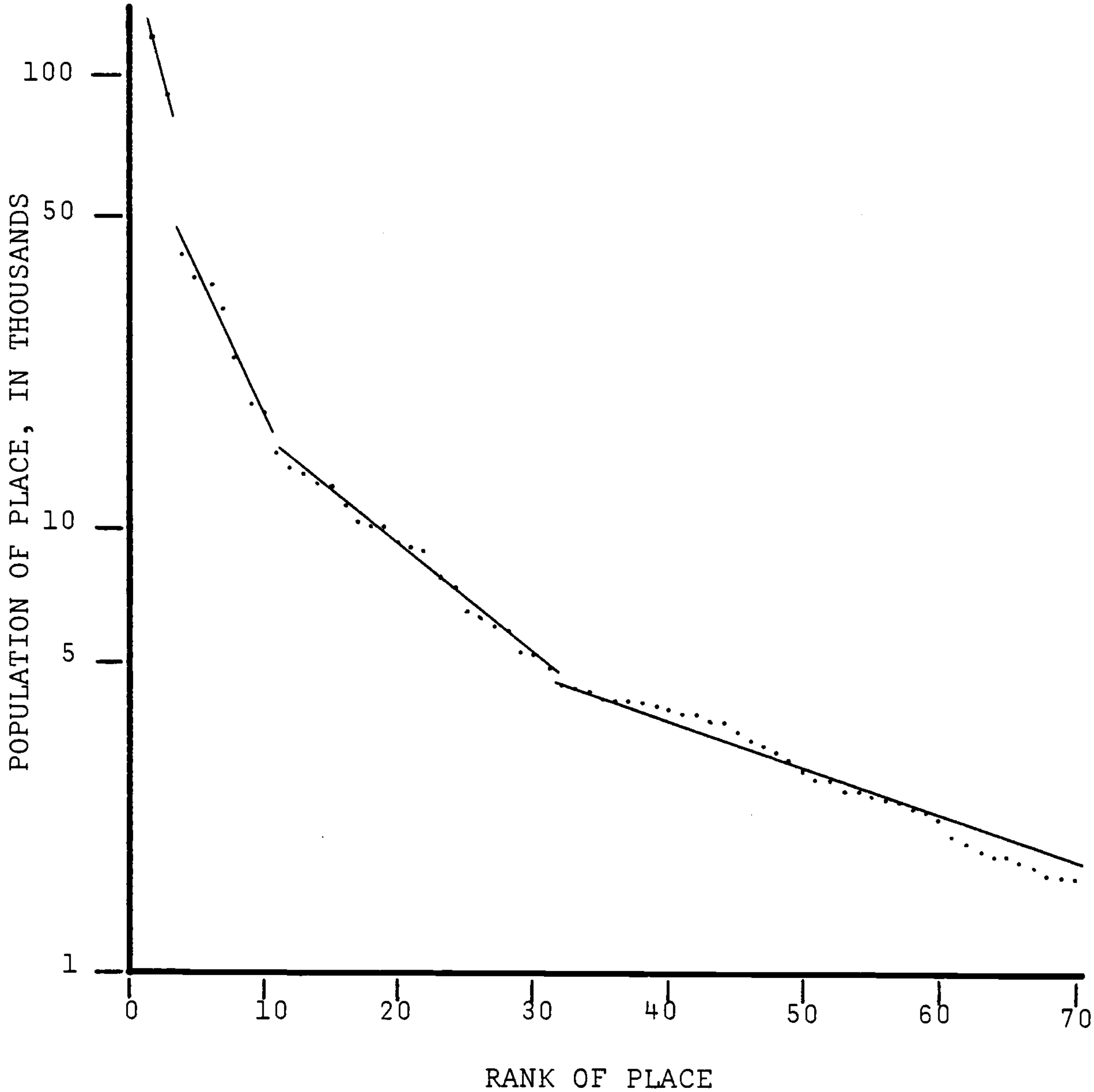


Fig. 3. Representative semi-log. plot of urban rank-size relationships with fitted regression equations: Portland, Oregon region, 1970.

both sides of apparent breaks in slope or at gaps in these data. All of the regressions yielded results indicating significant linear approximations. However, in order to select a single set of line segments as representative of the rank size curve and indicative of the central place associations for each data set a more qualitative selection process was conducted. This selection process consisted of comparing the regression equations for each line segment of each data set and comparing combinations of equations as a model of the whole system. Frequently, a line segment would stand out with obvious linearity and obvious starting and ending points. In these cases, the adjacent lines were selected to fit the obvious segment. In other cases, the segments would be selected so that they would coincide with a relatively large gap or a break in slope.

As an example of this technique, consider the 1970 data set (Figure 3). Because of its large size and regional capital character, Portland is set apart as a first order place in this regional system. The second and third ranked places, Eugene and Salem respectively, also stand apart from the rest of these data and are classified as second order places. The line containing the third order places starts with the fourth ranked city, (since the first three are already classified), and continues through the eleventh ranked city. However, the eleventh ranked city is separated from the tenth ranked city by a relatively large gap. Since the regression line is not significantly affected by the exclusion of the eleventh ranked place, and since exclusion of the tenth, ninth, or eighth ranked places does affect the regression equation, we can tenatively conclude that the eleventh ranked place is not a member of the third order. This exclusion could later be justified by comparing the eleventh ranked city to members of the third and fourth orders. Mean population comparisons as well as other qualitative and economic considerations would support its exclusion. A similar procedure is used to select the members of the other orders. The members of the fifth order were estimated by regressing the number of members in each order against its corresponding order for each data set (Appendices K and L). The resulting equation was then used to estimate the number of members of the next order. Since the population of members of this order extended below 1,000 inhabitants, this technique was not repeated for other orders because of the aforementioned unreliability of census data below this limit.

Semi-logarithmic Rank Size Regressions

All of the line segments finally selected yielded significant linear regressions and most of these lines also produced the best linear fit (Appendix G). An examination of the regression equations yielded several interesting observations: first, the slope parameters for the equations of each order were statistically similar; second, these slope parameters were statistically unlike the slope parameters of other order regressions; third, the slope decreased as the order and number of places increased; and fourth, the Y intercept, ln a, tended to increase through time. These observations led to a series of speculations about their significance.

When full logarithmic data are transferred to semi-logarithmic paper, straight lines become curved. These curved lines will resemble the semi-logarithmic rank size graphs (Fig. 3). To determine the relationship between these two ways of expressing these same data, the subdivided orders created by the lines on the semi-logarithmic paper were analyzed as simple transformations of full logarithmic data and as separate full logarithmic subsets determined by the ordering technique. In order to test for similarity to the transformed full logarithmic data, it was necessary to create new data sets comprised of exact X and Y coordinates generated by the full logarithmic regression equations. If the data points thus obtained are subdivided into the same groupings as the original semi-logarithmic orders, we can compare the resulting regression equations on these separate orders to determine if the semi-logarithmic line segments accurately represent the full logarithmic plots from which they originated. If the data points in the original semi-logarithmic orders are plotted as full logarithmic data, we can compare regression equations again to determine which equation form is more accurate.

Using a t test, the slopes of the original and the theoretical regression equations were compared. The paired slopes, original versus theoretical, were found to be statistically equivalent for all but the fifth (estimated),order of places. This indicated that the line segments and correlated regression equations were accurate representations of the full logarithmic plots, as well as accurate approximations of the semi-logarithmic curves. If the points contained in each order are analyzed as full logarithmic regressions, the equation parameters should have approximated those of its source. When this was done, the resulting equations produced statistically dissimilar parameters and coefficients of determination consistently lower than those of the corresponding semi-logarithmic equations, except for the fifth (estimated) orders. This indicated that the full logarithmic rank size uniformity evident in the whole data sets did not characterize the data set when it was subdivided. The rank size rule actually concealed more information about the internal structure of these data than it revealed. It appears then that the distribution of cities within this region is more accurately described by a hierarchical series of characteristic semi-logarithmic or exponential equations than by classic rank size rule.

Since these central place order regression equations described subsets of data described by a full logarithmic or geometric equation, this relationship would be reflected in the parameters of the regresssion equations, specifically in the slope parameters and in the number of elements per order. When these two parameters were plotted separately on semi-logarithmic paper for each data set, the expected linearity emerged. Semi-logarithmic regressions of these data confirmed the linear distribution (Appendices H through L). The slope parameter of the slope-order regressions all approximated one, again indicating a close approximation of the geometric distribution that characterized the whole data set.

Central Place Characteristics:Slopes

The internal ordering of the geometric rank size distribution as revealed by the semi-logarithmic analysis may correspond to the central place structure of the region. This expectation is supported by the internal consistency of these data. This apparent central place ordering is not distinctly step-like as Christaller theorized. However, the statistical comparisons among and between the various order regressions from the different data sets confirmed the intra-order similarities and the inter-order differences. Although each of these internal orders has characteristics unique to its order as Christaller claimed, the distribution of population within each order is exponential rather than identical. The consistency of the order regressions slopes for each order may be indicative of an equilibrium condition characteristic to each order, like that claimed by Zipf in his balance of the forces of diversification and concentration. This would indicate that places in these orders are subject to forces characteristic of their roles in the urban hierarchy. The lower slopes of the higher orders (smaller places) could be a reflection of the reliance of these places on low trade volume and their dependence on the native resource and labor base. The steeper slopes of the higher order places (larger cities) could be a reflection of their higher trade volume, diverse resource and labor base, their greater ability to generate consumption within their trade area, and perhaps their reliance on finished and semi-finished materials in their manufacturing. It would follow that other phenomena may be characteristically associated with each order in a similar manner. Although these generalizations are consistent with both the rank size rule and Zipf's Principle of Least Effort, and Christaller's description of the functional basis of central place theory, the urban structure described by this model consists of cities exponentially ordered into an exponential hierarchy of orders.

If we accept the model and theoretical structure outlined thus far, the order regression equations can be used as a base for historical analysis. We expect that in an urbanized region the higher orders would remain relatively stable because urban growth usually occurs in the larger urban areas rather than the smaller towns and villages; also, the number and spatial distribution of higher order places tends to balance deviations in the regression model. The urban order hierarchy that has been constructed for 1940 revealed a gap between the first and third orders (Appendices H through L). This gap was indicative of the relative primacy of Portland during this period, a role not revealed by B. Berry's primacy technique, but one that was apparent in the relative differences between the population of Portland and the second ranked place. In the 1940 to 1960 period, Portland was ten, eight, and six times the size of the second ranked place, respectively. During the 1940 and 1950 period, an eighty-five percent increase was recorded for cities in the 10,000 to 50,000 population range. This filling in of the second and third orders was also reflected in their slopes in the 1950 order regressions.

Central Place Characteristics:Orders and Elements

When the orders were arranged in increasing order the number of places contained in each order exhibited the exponential regularity observed by Christaller (Appendices J and K). If these values are taken as a pyramid of numbers, their distributions approximate Christaller's K=3 network. The actual ratio of number in each order to the number in each succeeding order can be found by taking the anti-logarithm of the slope parameter of the regression equations (Appendix L). The exact values for the 1940 and 1970 data sets are 2.64, 2.65, 2.82 and 2.83, respectively. This technique has also been employed by A. N. Strahler to compute the bifrocation ratio of streams.[14] This ratio has been used to characterize and compare different river basins. Similarly, Christaller uses this term, as in the K=3, 4, or 7 networks, to characterize urban systems. However, he did not use this value as a tool for comparing different regional urban systems to each other, he used it only to compare them to the ideal K=3, 4, or 7 networks. Since this term is a ratio, and is easily computed by linear regression techniques, it could be used as an index of urbanization and to facilitate interregional comparisons of urban systems.

From the data on the study area, I was able to define only five orders in the places of more than 1,000 population. Christaller defined seven orders within this same range of population. However, analysis of data from the study area indicated that comparisons based on population data may be unreliable. The Y intercept, lna, parameter, showed a consistent tendency to increase in the order regressions (Appendix G). This was also reflected in the means for each of the orders (Appendices M and N). This tendency to increase in mean population and Y intercept through time in the study area could be explained as the result of continual growth of the population in the face of a relatively constant urban to rural ratio. Certainly if direct population comparisons were subject to error within this study area and study period, comparisons between or within regions not subject to the same controls would be unpredictable.

With this caveat, I compared the results of the central place analysis of the study area with some of the central place generalizations made in two other studies.

<u>Comparison with Two Central Place Studies</u>

In his 1953 paper on the central places of southwestern Wisconsin, J. Brush employed a semi-logarithmic rank size plot to determine possible central place orders. He was successful with this technique, probably because his approximately square study area was not a coherent region and did not include an adequate functional system of cities.[15] Using other measures of centrality, he developed a list of the population and services that characterized hamlets, villages, and towns. One function he associated with the central places he called towns was that of county seat. County seats were significantly associated with the fourth order places in the present study, indicating a correlation between these fourth order places and Brush's "towns", (Appendix O). In his 1971 paper on the central place structure of the Pacific Northwest, from which the study area for the present paper was selected, R. Preston also revealed five orders in the Portland region.[16] His analysis utilized a more complex procedure and he apparently did not use agglomerated urban area data. Thus, a direct comparison of his results with those of the present study is not possible. Nevertheless, comparing the 1960 data, the technique presented in this paper correctly identified each of the members of his first three central place orders, 10 of 20 places in his fourth order, and 13 or his 34 fifth order places.

This limited comparison does not thoroughly validate this technique or provide unequivocal proof of the accuracy of its results. In order to do so, these results would have to be compared to those of other accepted central place techniques. Nevertheless, the application of this technique in the study region yielded reasonable representations of the central place system apparent in the region. Furthermore, the results were produced within the context of the generally accepted observations of the rank size rule and central place theory.

SUMMARY

The model of urban structure for the Portland, Oregon region during the 1940 to 1970 period was developed using only census population data and a technique employing linear regressions. These population data were aggregated to reflect the natural rather than the political boundaries of the urban places whenever possible. The study area selected was previously defined as a coherent region with a functional system of cities. The rank size distribution of cities within this region was accurately described by a straight line on full logarithmic graph paper, and by a geometric equation. When these data were transferred to semi-logarithmic graph paper, with rank on the arithmetic scale, they revealed a curved line apparently composed of linear segments with slopes decreasing with increasing rank. These data contained within each of these apparent linear segments were more accurately described by semi-logarithmic, or exponential, rather than logarithmic, or geometric, distributions and equations. The regression equations of these ordered line segments revealed qualities characteristic of central place theory. When the number of places in each order was regressed against its order, the slope parameter was found to approximate Christaller's K=3 network. This value could be used as an index of urbanization and as a comparative value for interregional comparisons of urban structure. The individual slope parameters of the segment,

or order, regression equations were found to be characteristic of that order, useful in the analysis of historical trends in urbanization, and as possible indicators of a characteristic equilibrium condition for each order. The rank size and central place conclusions revealed by this technique are consistent with similar research in this and in other regions. Although the model of urban structure revealed by this technique is consistent with the rank size rule, the logarithmic distribution was found to conceal the internal structure of the region. Instead, this region was more accurately described by an exponential hierarchy of orders. Although these internal orders were consistent with central place theory, the population of cities within each order were not found to be nearly equal as the classical theory claims but to be related to each other in a characteristic exponential fashion.

FOOTNOTES

1. G. K. Zipf, Human Behavior and the Principle of Least Effort facsimile of 1949 edition (New York: Hafner, 1965), pp. 348-387.

2. C. W. Baskin, A Critique and Translation of Walter Christaller's "Die Zentralen Orte in Süddeutschland" (Ann Arbor: University Microfilms, 1961).

3. B. J. L. Berry and A. Pred, Central Place Studies: A Bibliography of Theory and Applications. bibliography series no. 1, with supplement, of the Regional Science Research Institute, G. P. O. Box 8776, Philadelphia, PA 19101, pp. 3-11, pp. 15-18.

   B. J. L. Berry and W. L. Garrison, "Alternate Explanations of Urban Rank-Size Relationships," Annals, Association of American Geographers, Vol. 48 (1958), pp. 83-91, as well as Zipf, op. cit., footnote 2.

4. W. J. Porton and G. S. Lumis, "Some Graphs and Their Functional Forms," Technical Report No. 153, July 1972 (U.S. International Biological Program, Grassland Biome, 1972).

5. B. J. L. Berry and W. L. Garrison, op. cit., footnote 3, also, C. W. Baskin, op. cit., footnote 2.

6. C. W. Baskin, op. cit., footnote 2, pp. 21-23, p. 65, pp. 202-220, and p. 352.

7. Most notable is the criticism of R. Vining, "A Description of Certain Spatial Aspects of an Economic System," Economic Development and Cultural Change, Vol. 3 (1954-1955), pp. 147-195, also. B. J. L. Berry and W. L. Garrison, op. cit., footnote 3, and B. J. L. Berry, "Cities as Systems within Systems of Cities," Papers, Regional Science Association, Vol. 13 (1964), pp. 147-164.

8. B. J. L. Berry and W. L. Garrison, op. cit., footnote 3, B. J. L. Berry, op. cit., footnote 7, B. J. L. Berry, "City Size Distribution and Economic Development," Economic Development and Cultural Change, Vol. 9 (1961), pp. 573-588, and E. C. R. Reeve and J. S. Huxley, "Some Problems in the Study of Allometric Growth," in L. G. Clark ed., Essays on Growth and Form (New York: Harper, 1958), pp. 121-156.

9. C. W. Baskin, op. cit., footnote 1, pp. 1-7.

10. R. E. Preston, "The Structure of Central Place Systems," Economic Geography, Vol. 47 (1971), pp. 136-155.

11. W. Warntz, "A New Map of the Surface of Population Potentials for the United States, 1960," Geographical Review, Vol. 54 (1964), pp. 170-184, see W. Isard, Methods of Regional Analysis: an Introduction to Regional Science (Cambridge: MIT Press, 1960), for a discussion of gravity models.

12. The technique discussed in the remainder of this paper was developed to indicate regional structure, by which is meant, urban population distribution, hierarchy, orders of places, and members of those orders. Data sources are provided in the decennial census, U.S. Bureau of the Census, U.S. Census of Population: Number of Inhabitants, Oregon and in the same series, Washington, for the years 1940, 1950, 1960, and 1970.

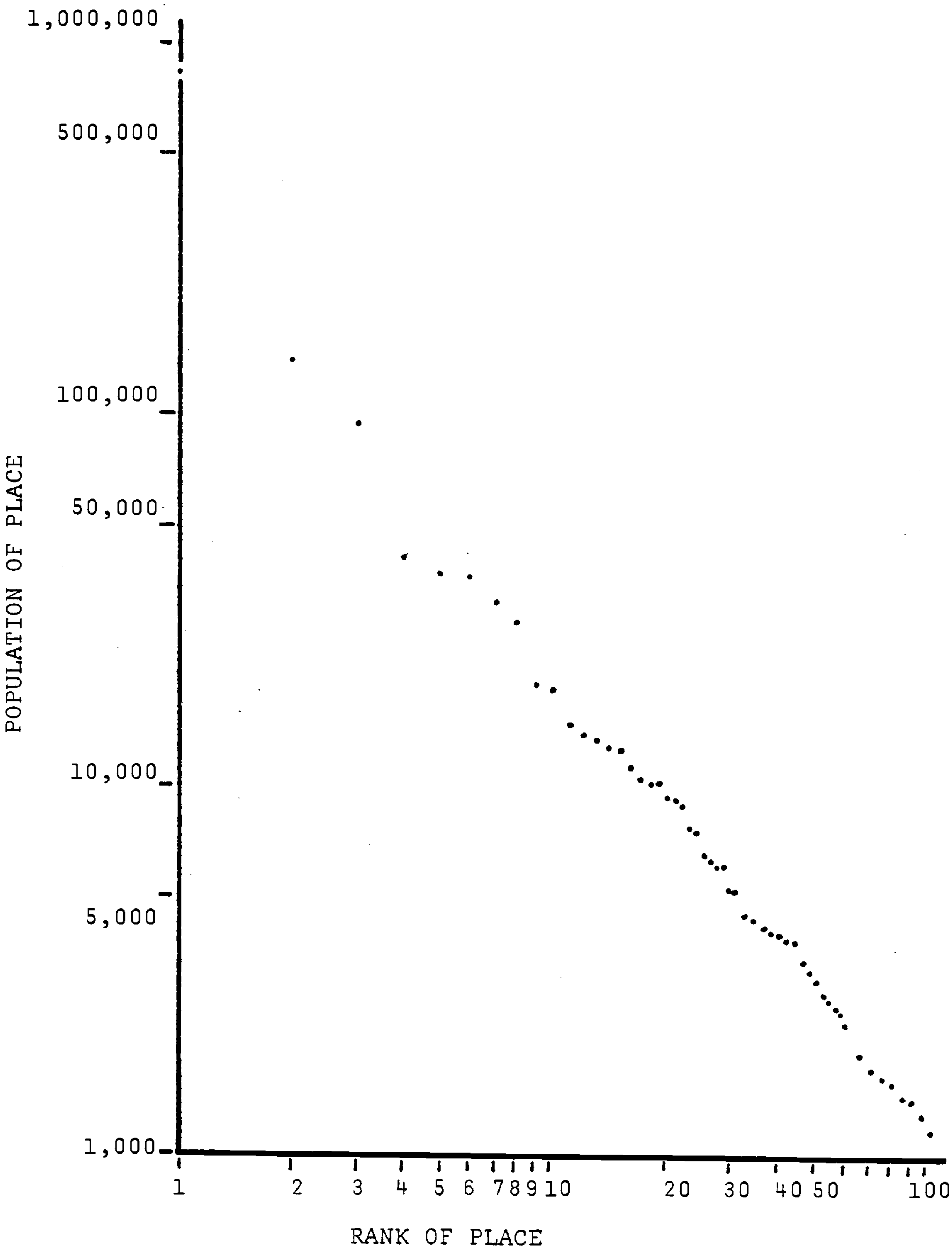
13. B. J. L. Berry, op. cit., footnote 8, "City Size Distribution . . .".

14. R. E. Horton, "Erosional Development of Streams and Their Drainage Basins: Hydrophysical Approach to Quantitative Morphology," Bulletin of the Geological Society of America, Vol. 56 (1945), pp. 275-306, and A. N. Strahler, "Quantitative Analysis of Watershed Geomorphology," Transactions, of the American Geophysical Union, Vol. 38 (1957), pp. 913-918, and A. N. Strahler, "Quantitative Geomorphology of Drainage Basins and Channel Networks," in Ven Te Chow (Ed.), Handbook of Applied Hydrology.

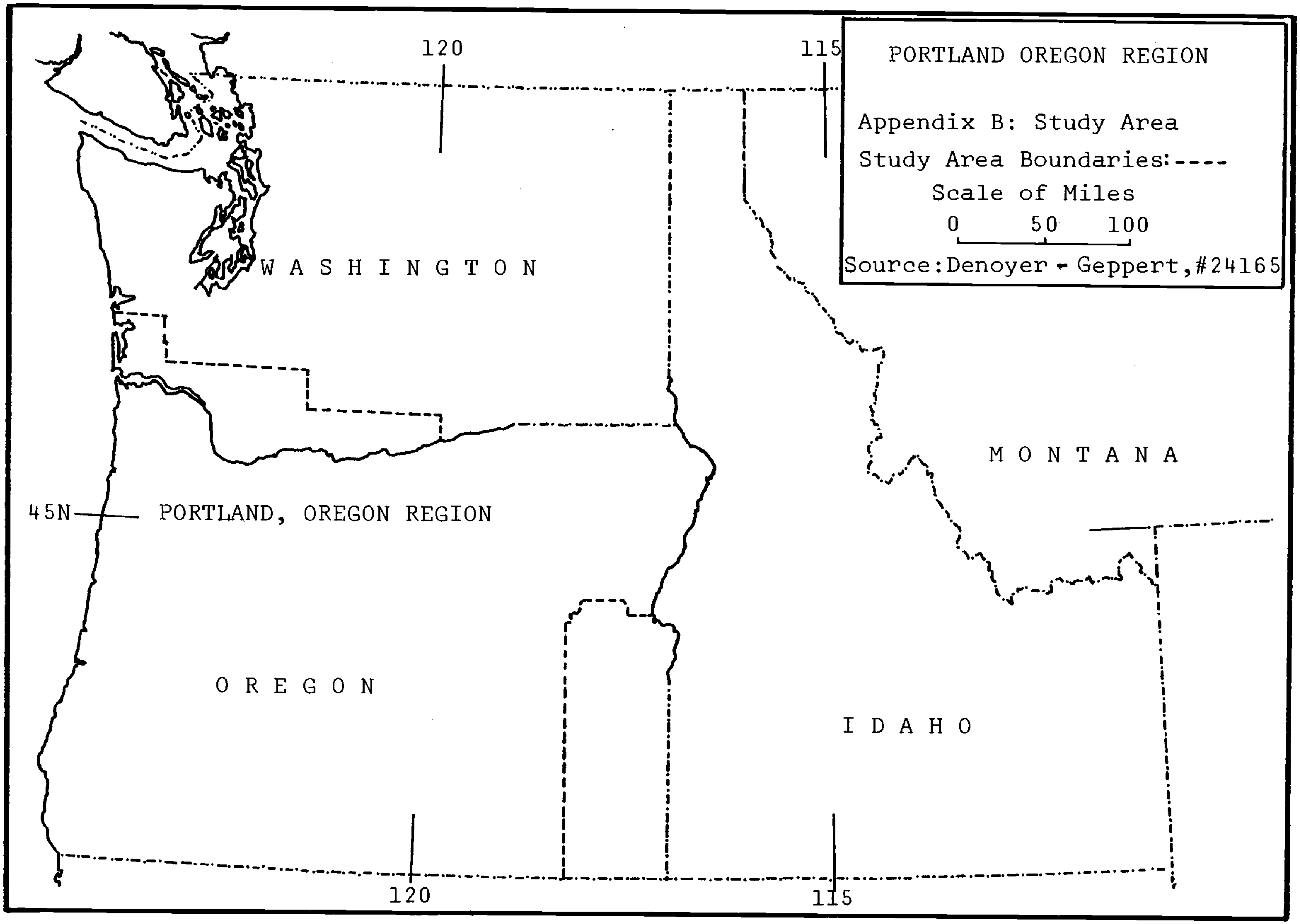
15. J. E. Brush, "The Hierarchy of Central Places in Southwestern Wisconsin," Geographical Review, Vol. 43 (1953), pp. 380-402, interestingly he did not refer to the semi-log. plotting technique in his later recapitulation of his research in J. E. Brush and H. E. Brocey, "Rural Service Centers in Southwestern Wisconsin and Southern England," Geographical Review, Vol. 45 (1955), pp. 559-569.

16. R. E. Preston, op. cit., footnote 10.

APPENDICES



Appendix A. Representative log.-log. plot of urban rank-size relationship: Portland, Oregon region, 1970


PORTLAND OREGON REGION
Appendix B: Study Area
Study Area Boundaries: ----
Scale of Miles
0 50 100
Source: Denoyer - Geppert, #24165
120
115
WASHINGTON
MONTANA
45N
PORTLAND, OREGON REGION
OREGON
IDAHO
120
115

Appendix C. Log.-log. regressions of the 100 largest places in the Portland, Oregon Region: 1940-1970

Equation $Y = a\,X^{b}$

Linear Regression Equation $\ln Y = \ln a + \ln X \cdot b$

Model	ln a	b, slope parameter	r^2
1940	12.027	-1.2365	.98403215
1950	12.272	-1.2040	.98621363
1960	12.534	-1.2258	.98791297
1970	12.803	-1.2484	.98927259

Appendix D. Log.-log. regressions of 99 largest places in the Portland, Oregon Region, excluding Portland: 1940-1970

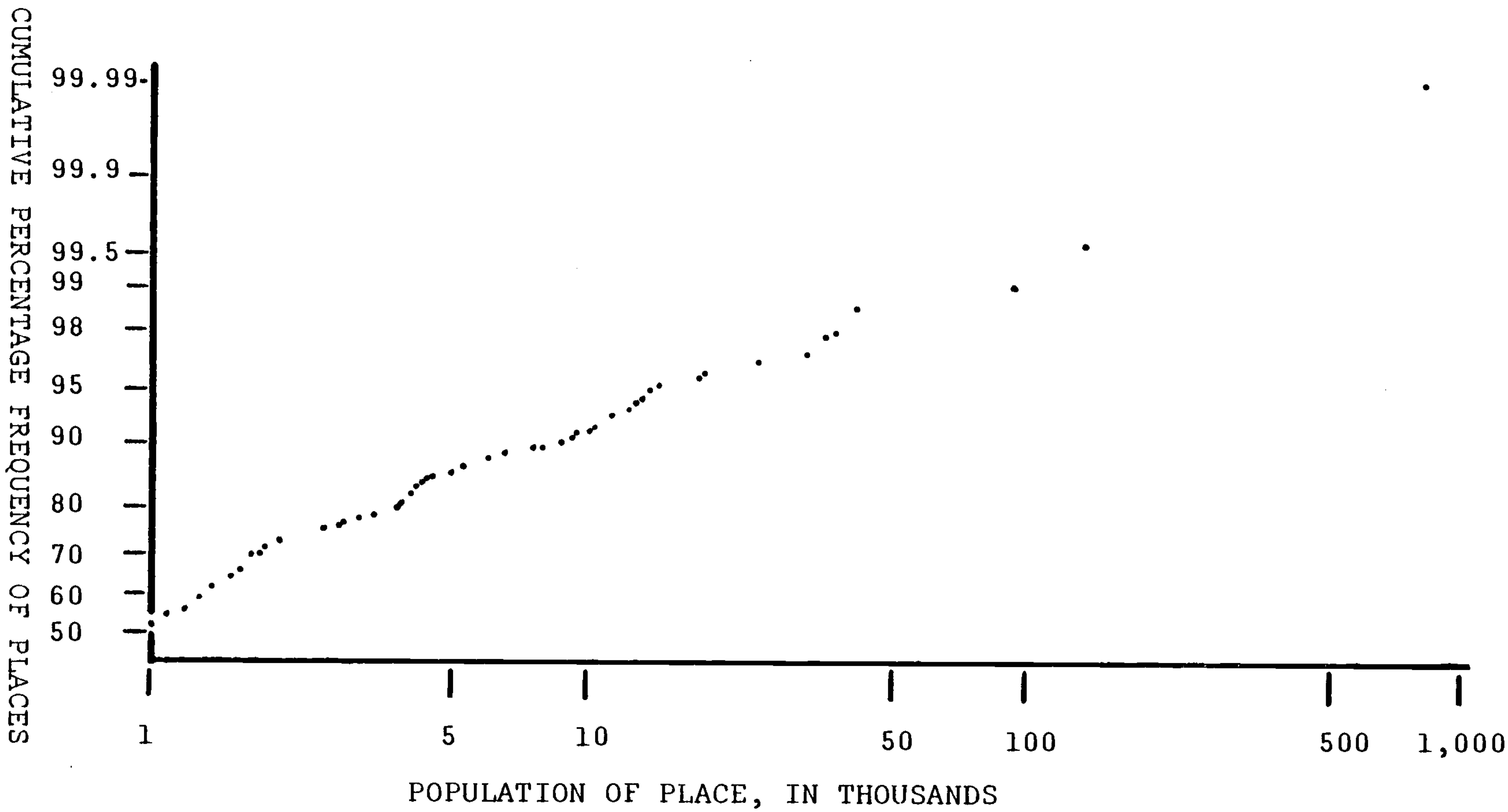
Equation $Y = a X^{b}$

Linear Regression Equation $\ln Y = \ln a + Z \cdot b$

Model	ln a	b, slope parameter	r^2
1940	11.851	-1.1911	.98865985
1950	12.099	-1.1593	.99168677
1960	12.365	-1.1821	.99300877
1970	12.641	-1.2065	.99396754

Appendix E. Predicted and actual populations of Portland, Oregon: 1940-1970

Model	Appendix C Regression Model	Appendix D Regression Model	Actual Population
1940	167,209	140,225	406,406
1950	213,630	179,692	512,643
1960	277,618	234,451	651,685
1970	363,306	308,970	824,926



Appendix F: Representative cumulative frequency vs. log. plot of primacy relationships: Portland, Oregon region, 1970.

Appendix G. Semi-log. regressions of apparent central place orders

Exponential Equation $Y = ae^{bX}$

Linear Regression Equation $\ln Y = \ln a + b \cdot X$

Class	Model	n	ln a	b slope parameter	r^2
2	1950	2	11.376	-.19533	1.0
2	1960	2	12.281	-.40624	1.0
2	1970	2	12.651	-.40327	1.0
3	1940	6	10.996	-.23369	.97601131
3	1950	5	10.809	-.14751	.91163869
3	1960	6	11.096	-.15859	.93580325
3	1970	7	11.287	-.14881	.94151671
4	1940	19	9.8044	-.069328	.98291143
4	1950	19	10.060	-.066554	.99179714
4	1960	22	10.159	-.061333	.98002914
4	1970	21	10.227	-.055562	.98763771
5, est.	1940	46	8.8203	-.031101	.96609366
5, est.	1950	42	9.1127	-.029712	.96823928
5, est.	1960	51	9.0044	-.023709	.9909856
5, est.	1970	55	9.1686	-.023832	.96980493

Appendix H: Semi-log. Plots of Regressions Slopes vs. Order: 1940-1970.

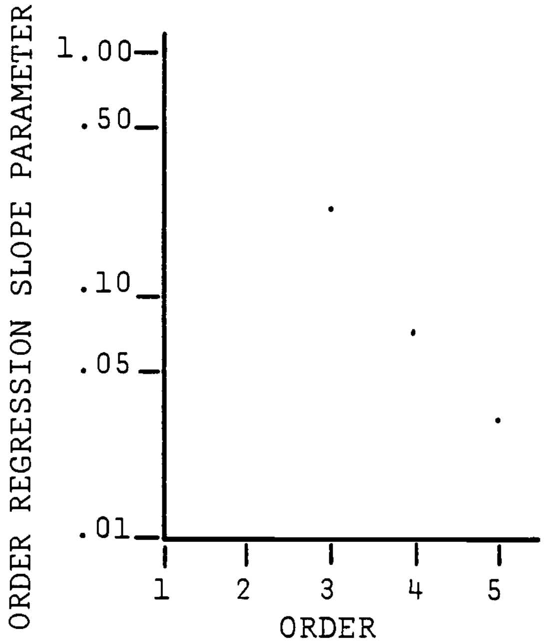


Fig. H1. Semi-log. plot of order regression slopes against order: Portland, Oregon region, 1940

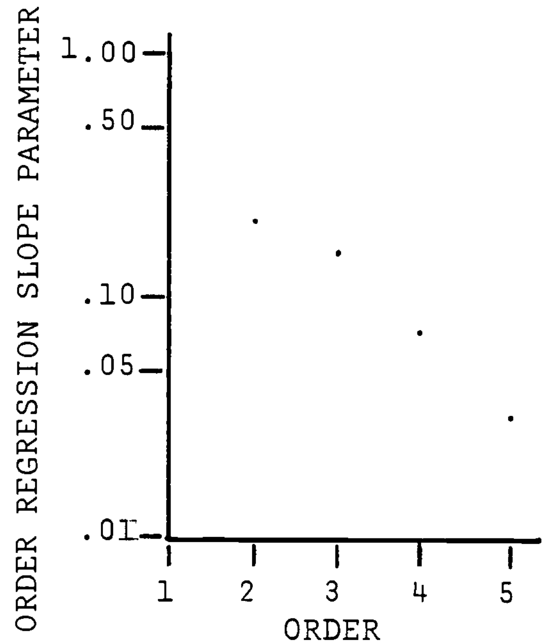


Fig. H2. Semi-log. plot of order regression slopes against order: Portland, Oregon region, 1950

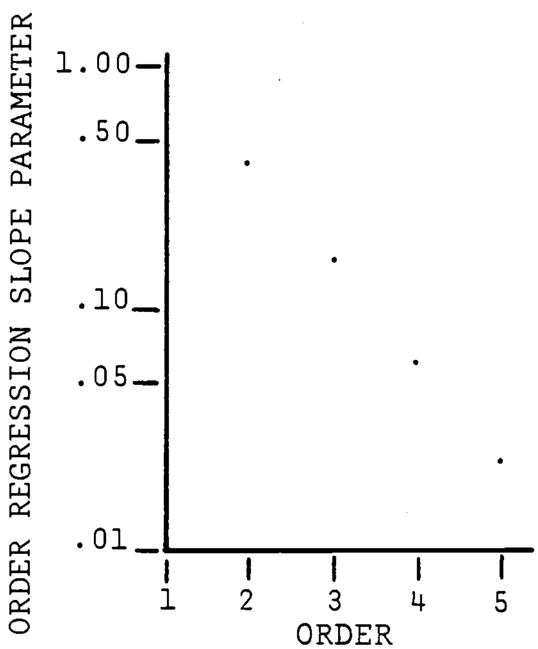


Fig. H3. Semi-log. plot of order regression slopes against order: Portland, Oregon region, 1960

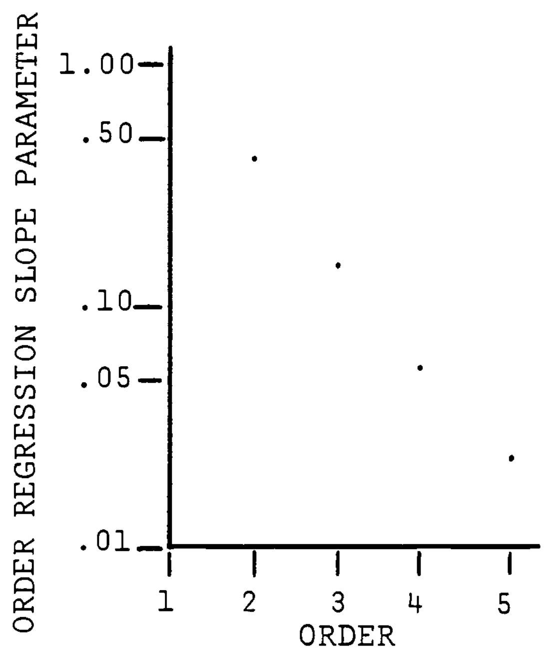


Fig. H4. Semi-log. plot of order regression slopes against order: Portland, Oregon region, 1970

Appendix I. Regressions of order slope parameters vs. order

Exponential Equation $Y = ae^{bX}$				
Linear Regression Equation $\ln Y - \ln a + b \cdot X$				
Model	n	ln a	b, slope parameter	r^2
1940	3	1.502450	-1.008378	.986178
1950	3	.486304	- .801172	.999947
1960	4	.996750	- .947327	.999994
1970	4	.954657	- .946986	.998586

Appendix J. Number of places for each order

Year	Order				
	1	2	3	4	5 (estimated)
1940	1	0	6	19	46
1950	1	2	5	19	42
1960	1	2	6	22	51
1970	1	2	7	21	55

Appendix K: Semi-log. Plots of Cities/Order vs. Order

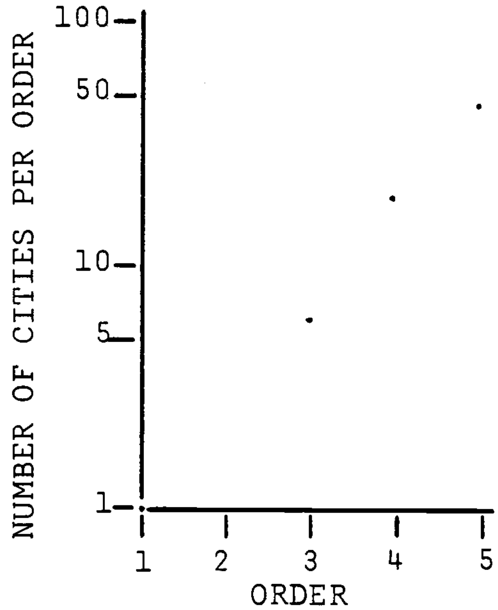


Fig. K1. Semi-log. plot of places per urban order against order: Portland, Oregon region, 1940.

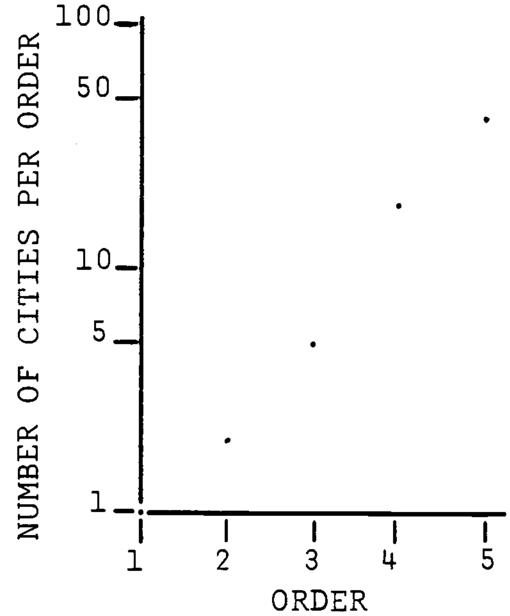


Fig. K2. Semi-log. plot of places per urban order against order: Portland, Oregon region, 1950.

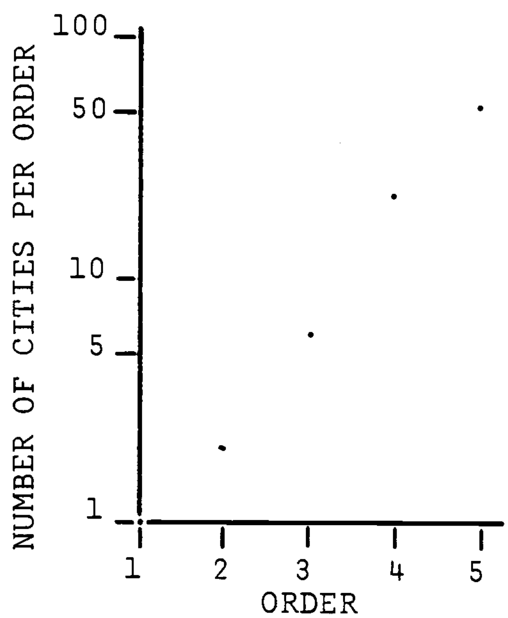


Fig. K3. Semi-log. plot of places per urban order against order: Portland, Oregon region, 1960.

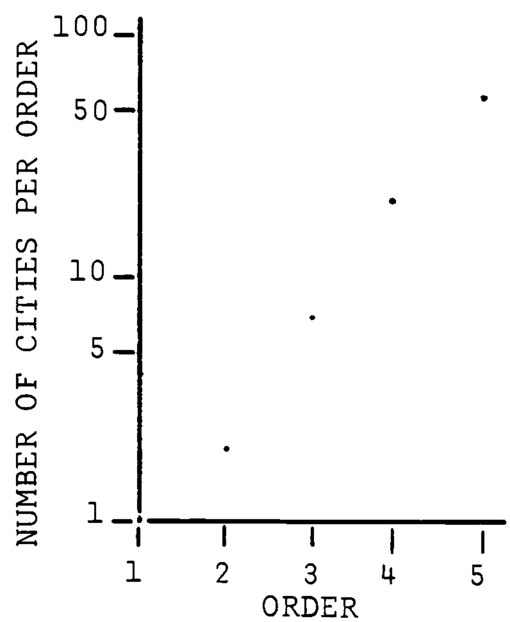


Fig. K4. Semi-log. plot of places per urban order against order: Portland, Oregon region, 1970.

Appendix L. Regressions of places per order vs. order

Exponential Equation $Y = ae^{bX}$

Linear Regression Equation $\ln Y = \ln a + b \cdot X$

Model	n	ln a	b slope parameter	r^2
1940	3	-1.0059	.96925	.99572063
1950	4	-1.1256	.97496	.97840387
1960	4	-1.1989	1.0372	.98283316
1970	4	-1.1757	1.0386	.98780785

Appendix M. Arithmetic mean population for each order

Order:	2	3	4	5 (estimated)
1940	NA	22,622	5,983	1,611
1950	53,747	20,920	7,538	2,312
1960	79,714	26,410	8,197	2,241
1970	116,148	29,447	9,103	2,524

Appendix N. Geometric mean population for each order

Order:	2	3	4	5 (estimated)
1940	NA	20,844	5,574	1,452
1950	53,491	20,420	7,062	2,147
1960	78,697	23,511	7,344	2,107
1970	113,826	28,155	8,610	2,351

Appendix O. County seats per number of places for each order

Order:	1	2	3	4	5 (estimated)	6 (remainder)
1940	3/1	NA	5/6	16/19	10/46	7/124
1950	3/1	2/2	4/5	13/19	6/42	8/144
1960	3/1	2/2	5/6	13/23	10/51	6/148
1970	3/1	2/2	5/7	11/21	12/55	6/178