THE PROCESS OF URBANIZATION: A COMPARATIVE STUDY WITH SPECIAL REFERENCE TO THE UNITED STATES, JAPAN AND INDIA: 1890-1950

by

MARK LAWNER

B.A., Harvard University, 1960

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF CITY PLANNING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June, 1963

Signature of Author..
Department of City Planning, May 17, 1963

Certified by..
Thesis Supervisor

Accepted by..
Chairman, Departmental Committee on Theses

-Dedication-

To Harry, Irene, Lynne and Robert

Table of Contents

.....there is no such thing as "proper territorial limits" of a community; but only approximate limits which are constantly shifting. A national population is continuous in its spread over its national territory, varying in density in what we call "communities"; these "communities" are not discrete entities, but make sense only in reference to the rest of the country. Yet we have found a useful thing to measure in these "communities" and by first considering them as discontinuous entities, we shall perhaps later be better able to comprehend them as they appear in the continuous cobweb of interrelated human organization.

-George K. Zipf

Acknowledgement

The author wishes to thank his advisor, Professor John Friedmann, for reading and reviewing parts of the manuscript throughout the various stages of their completion. The author would also like to express his appreciation to Professors Aaron Fleisher and George Wadsworth for their review of the statistical aspects of the study. For the special application of the Lorenz curves in the text the author is indebted to Jim Watanabe of the Harvard School of Design. Finally, thanks are due to past and present classmates, Ross Harris, Bill Pokross and Paul Kolp, who have made suggestions about, and corrections of, the manuscript.

-Abstract-

The process of urbanization is a dynamic which is reflected in certain demographic outcomes. Therefore, a definition of this process is best couched in demographic terms. A few key variables and their permutations suffice to describe the process of urbanization for purposes of international comparison--total population, number of cities, city sizes, and city rates of growth. Six propositions derive from these variables, and they have been tested in the United States, Japan and India over the six decades between 1888 and 1950. Stewart's rule relating number of cities to the urban total population ratio is confirmed in all three countries. Tisdale's preconditions--increase in cities and in city size--are to be found in the United States and Japan but not in India except for the last two decades. While Zipf's rank-size rule does not accurately predict the city size distributions for the most part, considerable rank-order stability prevails among the individual cities in all three countries. Jefferson's "law of the primate city" predicts relative position in the United States and in India (with "dual primacy" as a qualification), but "primacy" does not increase steadily with time in any of the three countries. Finally Madden's generalization of retarding growth rates is confined to the U.S. experience; individual rates of growth have yet to peak in Japan and India. A description of the process of urbanization in the above terms is a necessary first step in the delineation of the performance characteristics of a "system of cities."

-Introduction-

> "The problem of political economy is posed not by the necessity of an individual making a choice under a given system of constraints but rather by the necessity of a society of people making a choice among alternative systems of constraints. For a legislative body to evaluate an existing system of constraints is for the members of that body to compare....the working properties of this system with the working properties of some altered form of the system. A legislative program is literally a system of constraints upon and prescriptions of the actions of individuals. The outcome of the adoption of one system rather than another is described in terms of limiting forms of distributions.[1]"

The system selected for study here encompasses the cities of a nation-state; the performance characteristics of these cities will be analyzed both in the aggregate and by selective size-class. The concept of a "system of cities" has been advanced by many writers, but the structure of such a system, not to mention its universality, is still in the process of being determined. We shall focus in this study on certain strategic demographic variables, e.g. number of cities, city sizes and decade rates of growth, total population, and the variety of ways in which they interrelate.

> .....whatever else may be entailed, there is general agreement that the factors which underlie any such system (of cities) are probably reflected in one way or another in the demographic...relations among cities. It is this general hypothesis that is the primary rationale for studying the relations.[2]

The above variables subsume numerous other demographic and economic variables such as natural increase, net migration and numbers and types of economic functions. Also they are considered independently of various other economic variables with which they may or may not correlate. The strategy employed here is to interrelate the number of cities, their population sizes and decade rates of growth as a first step in the delineation of a "system of cities". These demographic variables clearly do not exist in isolation; it is likely that they both partially reflect and condition functional relations among cities. But this aspect would take us well beyond the scope of this study. Consequently, the demographic variables will not be explicitly related to other, "outside" variables, but some use will be made of economic, historical and cultural factors in the interpretation of the study findings.

Various combinations of the four demographic variables will be shown below to define the process of urbanization, and the case upon which this definition rests will be put forward. The process of urbanization involves, above all, change over time, yet in the midst of this change, certain relationships persist. It is these parameters that we want to determine as well as the correlations of individual variables with each other; that is, this study will attempt to ascertain the underlying parameters of the process of urbanization in addition to establishing various linear correlations. Most of the empirical relationships to be tested in this study have been developed by writers who have generalized from the U.S. experience of urbanization. One of the main objects of the study will be to see if, and to what degree, these generalizations are applicable to India and Japan.

These two countries differ markedly from the U.S. in a number of important respects: among them land area, total population, degree of industrialization, cultural role of cities, historical development in general, etc. In view of such profound differences, it remains to be seen whether or not their urbanization has been similar to that of the U.S. Except for the work of Kingsley Davis and his associates, comparative studies of urbanization are conspicuously absent from the literature. This is unfortunate, because the comparative approach allows one to study urbanization within a broad context and thus cull out the universal from the particular. The review of the existing literature will bring out the fact that research to date has focussed on problems of definition, the application of a single concept to one or more countries or the application of several concepts to a single country, but not the testing of a series of related or complementary propositions to several countries. The ideal, of course, would be to test a long series of interrelated propositions with reference to all the countries in the world. This study represents a modest beginning in that direction.

-Review of the Literature-

Hope Tisdale's definition of urbanization will serve as an entree into the literature, inasmuch as it is often cited by later writers on the subject; some build upon it, others try to substitute alternative formulations.

> "Urbanization is a process of population concentration. It proceeds in two ways: the multiplication of points of concentration and the increase in size of individual concentrations.
>
> Consistent with the definition of urbanization, cities may be defined as points of concentration. There is no need at this juncture to fix lower limits to the size and density which qualify a concentration as a city. There is no clear cut level of concentration at which a city suddenly springs into being. It is convenient from time to time to name certain levels beyond which concentrations are designated as cities. This is necessary in analyzing data and identifying characteristics of various size groups, but it does not alter the validity of the original concept.[3]"

She defends this kind of demographic definition on grounds that it a) avoids postulating the pre-existence of cities; b) disentangles causes and effects from urbanization itself; c) differentiates the end product, cities, from the process, urbanization; and d) avoids prejudging certain relationships such as the economic correlates of city size. Her definition has the added advantage of subsuming intervening variables such as natural increase and net migration.

To index an increase in the number and size of individual concentrations, Madden makes use of the mean size of all cities over 5,000 persons and also a modified Lorenz curve with percentage of urban population plotted on the vertical scale and percentage of urban places over 5,000 plotted on the horizontal scale. Thus increasing average city size is represented by an increase in the concavity of the resulting Lorenz curves.

Davis and Golden define urbanization in terms of the ratio of urban to total population. They agree with Tisdale that urbanization is not identical with cities, but differ from her in relating the growth of cities to total population instead of to previous city sizes and in ignoring the number and sizes of individual cities. Singer objects to the Davis-Golden formulation on the basis that their ratio obscures important differences such as those between one city of 500,000 and ten cities of 50,000 each. This defect diminishes the ratio's utility for comparative studies.

Stewart gets around this objection by explicitly relating the urban: total population ratio to the number of cities in the following empirical rule: $C = K \cdot U^x$ where C is the total number of cities, U the ratio of urban to total population and K and x are constants. Steward suspects that the values of K and x vary from country to country; his data show K equal to 10,450 and x equal to 2 in the U.S. This empirical rule represents a partial synthesis of the Eldridge and Davis-Golden formulations, but city size is still not taken fully into consideration.

The rank-size rule of Zipf fills this gap and parallels Stewart's rule, in that it relates the number of cities to city size directly, rather than by means of the urban to total population ratio; in effect, Zipf disaggregates the relationship. If the cities of a country are arrayed along a continuum by city size and rank-ordered accordingly, the relation of rank to size takes the form of a harmonic series, so that the rank of any given city times its size approximates a constant which in turn is equal to the size of the largest city in the country. In its simplest form the relation is expressed by the formula, $R_n \cdot S_n = M = S_1$, where R_n is the rank of the n^{th} city, S_n is the size of the n^{th} city, M is the constant and S_1 is the size of the largest city in the rank order.

> "However, since the rule holds only to a statistical approximation and is not rigorous, a better average fit in practice may be obtained by an adjustment of M to a value that is not exactly equal to the size of the largest city.[4]"

This latter formulation is known as the modified form of the rank-size rule, and it is represented by the formula $S = R^{-a} M$, which is identical to the simple form above, except that it allows the constant M to assume values other than the size of the largest city.

A number of alternative tests have been devised for the rank-size rule. Zipf employs a graphic technique whereby rank and size are plotted on log-log paper with linearity indicating correspondence to the rule. Zipf finds great linearity for time series in the U.S., Germany, Canada and France and for single years in India, Austro-Hungary and Hungary.

Singer anticipates Zipf by several years. He takes as a starting point Pareto's work on income data which rests on the same distribution. His test is both more precise and general than that of Zipf. It applies to size-classes rather than individual cities and it quantifies the degree of error. Singer uses the logarithmic form of the rank-size rule ($\log y = A + a \log x$) to predict the "expected" number of cities by size-class where y is the size class and x the number of cities that size or above. The constants, A and a, are derived from least-square lines

Table 1

COUNTRIES and their PARETO COEFFICIENTS

(a) in 19th and 20th CENTURY

DATES	Germany	France	U. S.	Sweden	Canada	England	Hungary	Spain	Japan
1871	1.45								
1872		1.20							
1876		1.24							
1880			1.11						
1881		1.23							
1885	1.25								
1886		1.21							
1890			1.08						
1891	1.19	1.20							
1895				.94					
1900			1.07						
1901					.98				
1910	1.15		1.09						
1911					1.05				
1920		1.30	1.03						1.59
1921					.96	.98			
1925	1.08								
1930				1.00			1.31	1.49	
1931					1.01				
1933	1.05								
1940								1.35	
1941					1.00		1.32		
1946	1.13	1.23							
1950			1.05	.99				1.17	1.16
1951					.98	.98			

Source: Singer, H.W., "The Courbes des Populations: A Parallel to Pareto's Law," Economic Journal, 46, June 1936, p. 259.

fit to the data for nine countries on twenty-three separate dates.

> "The inclination a of the straight line which in a logarithmic system of coordinates represents the relation between our x and y can be considered as an index of the relative frequency of small, medium and large towns which lends itself equally well to the purpose of international and historical comparison. The flatter this line, i.e. the smaller a , the greater is the proportion of large towns in a given number of towns.[5]"

The a values are called by some Pareto coefficients. The a values that Singer found are tabulated in Table 1. With those and the A constants Singer is able to substitute y into the logarithmic equation and generate "expected" numbers of cities in each size-class. The error in prediction is measured by the formula, $100 \frac{y-y'}{y}$, where y is the actual and y' the "expected" number of cities.

> "The arithmetical mean of these (errors) of the observations relating to the same country and the same time has been called the "average error." Therefore, it can be shown that the actual population distribution agrees with the law for different countries and different times in a very striking degree.[6]"

Allen extends this analysis to data for fifty-eight countries between 1946 and 1952; Table 2 summarizes his finding. For both conurbations and legal-political cities 27 of 44 countries have "average errors" under 7%, and only 9 of 44 are over 10%; for conurbations only, the figures become 15 of 21 and 2 of 21 respectively.

Table 2

Frequency Distribution

Number of countries where "average error" was:

		2.0-3.9	4.0-6.9	7.0-9.9	10.0 plus	Total
Europe	(a)	2	4	1	3	10
	(b)	2	4	1	1	8
Rest	(a)	2	4	3	4	13
	(b)	5	4	3	1	13
Total		11	16	8	9	44

(a) legal-political cities
(b) conurbations

Source: Allen, G.R., "The 'Courbes des Populations:' A Further Analysis," Bulletin of the Oxford University Institute of Statistics, 16, May-June 1954, p. 183.

"The main conclusion is, therefore, that the Pareto Law can be used with much success to summarize the relationships between size of towns and the number of towns at or above a specified size. Although the evidence does not show that it can be assumed as a universal law, it seems that given statistics relating to true conurbations the law would be much more forcefully supported.[7]"

Allen proceeds to comment on Singer's data for the Pareto coefficient:

......there has been a tendency to secure dispersions measured by Pareto coefficients of value between .90 and 1.10 (15 out of 21 countries) the coefficients of 19 out of 21 countries fall between .85 and 1.15.

With the exception of the U.S. the 19th century experienced a marked change in the degree of urbanization. For the beginning of the century Pareto coefficients have been obtained, or seem obtainable, of values greater than 1.40. The fall to stable values much nearer unity was completed by the end of the century.

(This) marked tendency for the Pareto coefficients for any one country to change very little in recent decades........is strongly at odds with the commonly held impression that in modern, western societies larger towns have been growing relative to the smaller towns.[8]

By comparison with the above, Hoyt, Stewart and Gibbs devise only partial tests of the rank-size rule. Hoyt takes five countries--Canada, Australia, United States, Russia, and China--and examines only the ten largest cities; in each case only several cities approximate their "expected" size. Then Hoyt also finds significant discontinuities in the upper end of the city size distribution in Japan, Germany, Russia and England.

Stewart makes note of similar discontinuities in his city size data. He compares the two largest cities in seventy-two countries in 1950 and compiles a frequency distribution for the ratio between the two:[9]

≠1 : ≠2 city	1-1.5	1.5-2	2-2.5	2.5-3	3-4	4-5	5-10	10+
≠ countries	14	6	9	2	20	3	16	2.

The ratio's median value is 3.2 versus an "expected" value of 2.0 according to the rank-size rule. But Stewart is actually not testing this rule at all but rather Mark Jefferson's "law of the primate city" which states that

> .....the largest city shall be supereminent, and not merely in size, but in national influence...once a city is larger than any other in its country, this mere fact gives it an impetus to grow that cannot affect any other city, and it draws away from all of them in character as well as in size....It becomes the primate city.
>
> In 28 of the leading countries of the world the largest city is more than twice as large as the next, in 18 more than three times as large--a constancy of recurrence that gives this relation the status of law. If we call the population of the largest city in each of these countries 100, the second city's number averages 30 and the third 20.[10]

His data are for the 1930's mostly; where data exist for two separate dates, primacy increases in 4 of 6 cases. "Nationalism," "general education" and "easy communications" are preconditions for primacy. The lack of primacy in English colonies is explained by the function of London as "primate" city for the Empire.

But other cases are not so easily explained. In Stewart's data the top city is less than twice the second city in twenty out of seventy-two countries (27.5%). Gibbs finds no "primacy" in Brazil (1950) or Italy (1951). In view of these exceptions, Jefferson's "law" is not a law at all but a generalization whose qualifications are not yet clear.

The question remains, does evidence of "primacy" constitute refutation of the rank-size rule. This writer does not think so for these reasons: the rank-size rule applies to the city size distribution not just a single pair of cities; the modified constant of the generalized rank-size rule reduces the amount of "primacy" in many cases; and a "primacy" of 2:1 is what would be "expected" by the rank-size rule.

Stewart claims that the rank size rule's

> ....applicability in the middle range (of city sizes) in large areas is due partly to diversity within the area in the values of the determinants of town size and spacing and partly to the gross coincidence of the rank-size curve and the town function pyramid. The rule is a better description for large, heterogeneous areas than for small homogeneous areas, where town size, spacing and function are most closely interconnected....Deviations from homogeneity tend to smooth the discrete size-classes into an approximation of the rank-size rule.[11]

But heterogeneity would seem to be the rule and not the exception in most countries. In addition cities do not fall into discrete size-classes automatically; the size-classes are delineated a priori by means of functional or statistical criteria. Vining observes that if cities are arrayed along an axis by size, they form a continuum. These cities may be ordered into size-classes only by the employment of some additional criterion. Thus he takes issue with those who hold that discrete size-classes are "naturally distinguishable <u>types</u> rather than mere expedients for describing phenomena that exhibit virtually continuous variation."[12]

Stewart's final objection to the rank-size rule is perhaps more valid. He argues that city size distributions assume an S-shape rather than a linear form, and advances supporting data from Sweden and Denmark. This hypothesis has yet to be tested systematically; such a test would require a comparison of predictions based on S-curves with those based on least-square lines.

Gibbs indexes "goodness" of fit of the rank-size rule by the "per cent of people who would have to move from one metropolitan area to another one to bring about complete conformity to the rule." This rather demanding test yields results of 8.2% and 13.2% for Brazil (1920, 1950), 9.0% and 9.3% for Canada (1921, 1956), 13.3% and 9.0% for Italy (1921, 1951) and 19.5% and 20.2% for France (1926, 1954).

> "Another aspect of change in the urban-size hierarchy is the movement of individual cities within the hierarchy. If the rates of growth for all cities during a period are the same, there is no change in their ranks in the hierarchy. Thus a comparison of the size ranks of cities at one point in time with that at a later time reveals how much movement has taken place during the period. A summary measure in this case is provided by a rank-order coefficient of correlation between the sise ranks of cities at two points in time.[13]"

For the same six countries and time periods as above, Gibbs' data reveals these rank-order correlations: +.81, +.87, +.97, +.97, +.88 and +.82; the country with least urban growth, France, also had the fewest rank shifts while the country with most urban growth, Brazil, experienced the greatest number of rank shifts. "Instability of city ranking suggests uneven economic development of different regions within a country. The major exception to this generalization is functionally specialized cities, such as manufacturing centers, which do not rely on the prosperity of the surrounding region."[14]

Madden is similarly much concerned with the rank shifts of individual cities over time. He remarks that while

> ....the size of a given city is not independent of the size of the largest city (and) the decade growth rate of a given city is not independent of the decade growth rate of other cities.....in no instance is the size or growth rate of any city absolutely determined. The regularity is statistical and implies a certain freedom for the individual city to move about within rather wide limits of size and growth rate.[15]

His data for the U.S. (1790-1950) very strikingly contrast the stability of the city size distribution as a whole with the considerable number of individual rank shifts. He delimits three size classes by the amount and direction of these city rank shifts over the time series: 14 cities shift less than 10 ranks, 22 shift between 10 and 100 ranks and 10 shift more than 100 ranks; 30 cities gain rank, 17 cities lose rank and 3 cities experienced no change. As for decade rates of growth, few cities grow at the average rate, and the variability of these individual rates appears to be associated to a high degree (+.91) with the average growth rate. The mean growth rates of both urban population and total population retard as the number of cities increases. Madden interprets this as heightened intercity competition for population and the substitution of "successful," i.e. growing, cities for certain other ones. Cities with population over 100,000 also exhibit a secular trend downward in their rate of growth; only 3 exceptions are found out of 103 cases. This retardation, however, takes place in a variety of ways. There appears to be little continuity in both growths and declines of the growth rates. Madden derives his index of continuity as follows. He designates negative decade to decade differences as minus one, positive differences as plus one and no change as zero; then he substracts the absolute number of minus ones from the absolute number of plus ones and divides by the total number of growth rates; the result is an index number ranging from -1.00 to +1.00. A value of plus one indicates a city has experienced a continuous increase in its rate of growth over six decades; zero means the city has witnessed as many downward shifts in growth rate as upward ones; a value of minus one shows that a city has been witnessing an increasingly rapid decline in its growth.

Vining interprets city size distributions somewhat differently than all the above writers. His view is that within a given land area there exists a human population system characterized by a certain amount of density peaks where population is relatively high and vast interstitial areas with low population densities. It is this density configuration and its dynamic that he takes as his central concern. His stochastic model will not be detailed here. Rather we will confine our review to his comments on the rank-size rule.

> "The fit of the simple formula...for the total are area holds only for relatively large values of R (total number of cities)....should there be a stability of form of the size distribution of cities...one may well guess that the frequency function specifying this stable form will be found to be complicated and involved and that the rank-size rule represents merely a rough approximation for large values of R.[16]

From our review of the literature it is clear that while urbanization is not a simple process, certain strategic demographic variables suffice to describe it for the purposes of international comparison. These variables are four in number--the number of cities, city sizes, city growth rates and total population--and they are both expressed and intercorrelated in a variety of statistical forms.

Briefly, the following relationships have been postulated in the literature. Stewart correlates the number of cities with the ratio of urban to total population by means of a constant factor; Tisdale links increases in number of cities to increases in city size (the two together yield an average city size); Zipf rank-orders cities by population size and then correlates these ranks with the city sizes themselves; Jefferson contrasts the size of the largest city with that of the second and third largest cities; Gibb compares successive rank-orders by city size and correlates their stability (or lack of it) with the decade rate of growth for all the cities in the rank order; Madden compares the mean decade growth rate for all cities with the standard deviation of city growth rates about that mean, and he contrasts the mean growth rates at the beginning of the time series with those at the end of it.

When the process of urbanization is described in these demographic terms, it becomes clearly distinguishable from its origins, its impact and concomitant processes. The strategic variables above do reflect numerous intervening variables, both demographic and economic, and thus they serve as convenient summary measures. Economy of expression is gained at the expense of empirical detail. Inasmuch as our main concern here is to study how urbanization varies from country to country, we must necessarily abstract to a high degree.

Comparative studies are conspicuous by their absence from the literature. The only exceptions to this are a few cases where a single relationship was tested in several countries or in a number of cities at one point in time. No systematic attempt to examine the process of urbanization, both as it varies from country to country and over a time series of any length, is to be found in the literature. This study represents an initial effort to fill that gap.

The United States, Japan and India are the three countries that have been selected for study, and the time series covers roughly two generations: 1890-1950 in the U.S., 1888-1950 in Japan and 1891-1951 in India. All of the propositions that will be tested in these three countries may be found in the summary paragraph above with three exceptions: rank stability (by city size) is correlated with increases in the number of cities; differences between mean growth rates (all cities) early and late in each time series are correlated with population increases in the average size city; and rank-orders by city size are compared with rank orders by city growth rate.

Propositions whose validity rests largely upon their confirmation in U.S. experience will be subjected to tests on the data from Japan and India, and the new formulations immediately above will meet their test in all three countries. Both international differences and similarities will need to be interpreted in the light of various fundamental differences among the three nations. Among these are land form, land area, total population, population increase, degree of industrialization, the cultural role of cities and historical development in general.

The topography of Japan is highly mountainous throughout most of the islands with a few, narrow alluvial plains between the mountains and the sea, while both the U.S. and India are large land masses with mountain ranges parallel to their east and west coastlines, some interior mountains and vast interior plains. As for size the U.S. is 7,839,000 square kilometers, India 4,675,000 and Japan only 382,000--a ratio of roughly 100-60-5 if the U.S. size is made equal to 100. In 1890 the U.S., Japanese and India total populations were 62,980,000 vs. 39,607,000 vs.235,900,000; in 1950 they were 151,326,000 vs. 83,200,000 vs. 356,900,000. Continuing the convention of equating the U.S. total with 100 we get the following two ratios: 100 vs. 63 vs. 375 and 100 vs. 55 vs. 236. Both India and Japan trace most of their population expansion to natural increase in contrast to the U.S. where in-migration has played a very large role.

In Rostow's terms and by his calculations, the relative dates for "takeoff" in the three economies are as follows: the U.S. 1843-1860, Japan 1878-1900 and India 1952- ?[17] This places Japan roughly forty years behind the U.S., and India an additional fifty to sixty years behind that. Rostow asserts that the U.S. completed its "drive to maturity" by 1900 whereas Japan reached that stage by 1940 leaving it still approximately forty years behind the U.S.; India is attempting one of the first planned "takeoffs" in economic history, and it remains to be seen when it will arrive at the stage of technological maturity.[18] By entering the stage of "mass consumption" in the postwar period, Japan has narrowed the gap between it and the U.S. to about thirty years.

The role of the Indian city differs from that of its counterparts in Japan and the U.S. There is a very strong rural tradition in India, and urban inhabitants return periodically to their native villages. This happened in the U.S. only during the Depression years. The large Indiancities are more on the order of conglomerations of "urban villages" than cities. Admittedly there is heterogeneity in the Japanese city where "traditional" and "modern" quarters exist side by side and in the U.S. where ethnic ghettoes are far from having disappeared, but their degree of isolation from each other appears to be less. There are some functional differences also important in this context. India has a large number of administrative and religious centers in its interior, outgrowths of the pre-British cities there. When the British arrived, they established port cities along both coasts which then served as entrepots and administrative centers. The Japanese are not without their port cities; the Pacific Ocean side of the islands contains four clusters of large, industrial cities where the alluvial plain is widest. Communications among these four have been either by railroad or by sea. The U.S. on the other hand has had good overland communications, and its development has displayed, among other aspects, a westward, urban movement and a distinct regional pattern of growth whereby new areas opened up and developed at distinctly separate points in time, e.g. the Northwest Territory, the Louisiana Purchase, the California gold rush, and so on.

This sketch of national differences includes a number of oversimplifications, but as the purpose in stating them is to establish rough benchmarks for the interpretation of the findings, they need not be too detailed.

Although their boundaries may change over time nation-states constitute relatively closed systems; population is substantially more mobile within countries than between them, and the rise of new cities takes place within national borders. Unfortunately, time precludes the application of our propositions to supra-national areas, whether whole continents, common market areas or the like; for the same reason, the impact of colonialism and post-colonialism on urbanization cannot be assessed.

The sixty-year time series data for the U.S., Japan and India are more or less comparable in several respects. First, there is a close coincidence among the relative census dates. The U.S. series begins in 1890, the Japanese series in 1888 and the Indian series in 1891. At no time in these series is there more than a three-year divergence of the relative census years. Because of this proximity the data have not been adjusted to a series of common dates. (There is one intercensus period which is not a decade but twelve years: Japan 1908-1920). The census data have been collected according to internally consistent definitions in each of the three countries through the sixty-year period. If and when the census definition has changed, a systematic bias has thereby been introduced which does not invalidate the use of these data for comparative purposes.

The use of data on legal-political cities also introduces a systematic bias, one that needs more detailed consideration. The bias encountered through city annexations is assumed to be negligible. While the legal-political city does tend to "underbound" the urbanized area in many cases, several strong arguments may be advanced in its behalf. The use of the legal-political city allows for comparability with earlier studies on the subject of urbanization. Political boundaries are not without some reality, a fact overlooked by most advocates of the metropolitan or urbanized area; city limits bound a decision-making unit that exercises a variety of controls on development from zoning to property taxes, subdivision controls and so on. The way in which these controls are exercised varies greatly from city to city. Further, data has been traditionally collected for legal-political cities making possible the compilation of reasonably length time series, whereas little data is currently available for continuously urbanized areas--one decade at best with a few exceptions. Finally, for comparative purposes the legal-political city may be as useful as the urbanized area. Gibbs has done the most systematic research on this matter to date. His

> ....results....strikingly belie the hypothesis that officially reported statistics on either urban agglomerations or administratively defined cities and towns are grossly non-comparable as between countries. The product-moment coefficients of correlation (r) between the percentage in any size-class of urban agglomerations...and the percentage in Metropolitan Areas is +.91 or above. As could be expected the correlations are generally lower between the percentages in cities and towns defined administratively and the proportion in Metropolitan Areas; but the difference between this and the first case <u>is amazingly small</u>.[19]

In the propositions that follow, a 5,000 person minimum size criterion is employed to define a city. This criterion is admittedly arbitrary; its use is dictated by the need to standardize the data so as to be able to make international comparisons. As long as no claims are advanced to the effect that one country is more urbanized than another by virtue of its having more cities over 5,000 or a greater percentage of total population in such cities, the usage of this minimum can be justified. Suburbs and satellite cities get included in the universe under study as soon as they cross the minimum threshold; therefore the populations of urbanized areas are not understated.

The size-class of cities over 100,000 is singled out for special attention in some of the propositions, largely for pragmatic reasons. In order to make detailed comparisons of the performance characteristics of individual cities in this study, their numbers have had to be reduced to manageable proportions. The 100,000 figure has proved to be a convenient cutoff point in that regard, and also in that the resulting size-class includes for the most part only those cities performing multiple functions.

Size-classes, however delineated, lend themselves to the analysis of time series data in two separate ways. The so-called "city" method involves the selection of a group of cities by size limits at some initial point and the subsequent tracing of their performance characteristics over the course of the time series; this method has the effect of holding the number of cities constant. On the other hand the "class" method entails the specification of a minimum size criterion at the initial point and the inclusion, at later points in time, of all cities rising above that minimum threshold. In this case, the number of cities is allowed to vary with time.

The propositions that are to be tested appear below. They derive largely from the existing literature on urbanization except for the additions enumerated above. The hope is that their range of application may be extended from the U.S. to Japan and India and that the relationships between propositions may become clearer. They have been stated in such an order as to facilitate the making of such connections;

the links that exist will be remarked upon as we go along. Each proposition is followed by what are considered to be appropriate measures to test its validity.

-The Propositions and Methodology for their Testing-

Proposition 1: The number of cities in a country varies with the ratio of urban to total population by a constant factor, so that at any one point in time one may use the value of either quantity to predict the other. To be more precise, $C = A \cdot U^{a}$, where C is the number of cities, U is the ratio of urban to total population, and A and a are empirically derived constants. To test this proposition, a least-squares line is fitted to the time series data with the number of cities plotted on the y-axis and the ratio of urban to total population on the x-axis. From these least-squares lines the values of A and a are obtained, A being the y-intercept and a the slope of the line. Into the logarithmic form of these least-squares lines, e.g. $\log C = A$ plus $a \cdot \log U$ are substituted the actual number of cities for each census data in the time series to generate successive, "expected" values for the ratio of urban to total population. These "expected" values are compared with the actual values of the ratio, and the standard error of estimate is computed which involves squaring the absolute error, summing these squares, dividing that sum by the total number of errors to get a mean and finally taking the square root of this mean error. The standard error of estimate allows one to say that 68% of the time, the error of our predictions can be expected to be less than some specified quantity, x .

Proposition 2: Increases in the average size of all cities are positively correlated with increases in the number of cities. Two measures of this association will be computed. The first is a Pearsonian product-moment correlation coefficient which serves to quantify the strength of the relationship between the two variables. The equation is

$r = \frac{S(xy)}{\sqrt{Sx^2 \cdot Sy^2}}$, where the x's are the differences between the actual average sizes of all cities and the mean for the group as a whole and the y's are the differences between the actual numbers of cities and the mean for the group as a whole. The correlation coefficients thus obtained will be squared and multiplied by 100 to establish the amount of variation of the y's that is attributable to their relationship with the x's . The second measure is a series of Lorenz curves whereby percent of total urban population is plotted on a y-axis and percent of total number of cities on the x-axis. These curves indicate the relative distribution of the aggregate (population) among its component parts (cities); the two extremes are 1) a straight line with slope of plus one and a y-intercept of 0.00 and 2) a straight line with slope of infinity and a y-intercept of 100.0. They represent situations where 1) all cities are the same size and 2) all the urban population is concentrated in one large city. Virtually all cases in reality fall somewhere between these two extremes. In this study we will make use of the y-intercept to index the degree of concentration; the y-intercept is determined by delineating tangent to the Lorenz curve a line whose slope is plus one and reading off the point where this tangent intercepts the y-axis. If the Lorenz curves become more concave with time (and thus the values of the y-intercepts increase), the proposition is confirmed. Historically this relation between number and size of cities tends to become subject to a kind of "law of diminishing return" during which the value of the average city size peaks and begins to retard. It remains to be seen whether the time series employed here is long enough to witness such a development.

Proposition 3: <u>The generalized rank-size rule in its logarithmic form,</u> $\log y = A + a (\log x)$, <u>provides a good prediction of the distribution of cities by size.</u> To test this proposition the procedures of Singer and Allen are followed. To a scatter diagram with greater-than size-classes on the x-axis and the cumulative number of cities by greater-than size-class on the y-axis (equivalent to rank), a least-square line is fit, and the values of the y-intercept and the slope, i.e. $\underline{A}$ and $\underline{a}$, are determined. These two parameters now given, the substitution of y-values (greater-than size-classes) into the logarithmic form of the rank-size rule yields x-values: the "expected" number of cities in each size-class. The "expected" number of cities are then compared with the "actual" number of cities, and the differential between the two is divided by the "actual" number of cities to establish the error of the prediction. Once the errors for each size class have been computed, they are summed and divided by the total number of size classes to get an "average error."

For each census date and for each country there is a unique least-square line with its y-intercept and slope (the Pareto coefficient). The slope has interesting properties somewhat independent of the rank-size rule itself, as does the x-intercept which represents the _M_ constant in the original formulation of the rank-size rule. The magnitude of the Pareto coefficient indexes the relative frequency of large cities vis-a-vis medium-sized and small cities; diminishing values indicate an increasing preponderance of large cities. This coefficient will be examined over time to determine whether there is a secular trend in its movement. The x-intercept, or _M_, is the hypothetical value that the largest city would have if there were a perfect correspondence of an actual city size distribution and the rank-size rule. The _M_ values will be compared with the actual sizes of the largest cities to see how close the approximation is. Finally, the two parameters, _A_ and _a_ , will be plotted against time on the same graph, _A_ along the y-axis and _a_ along the x-axis, to guage their variability over the course of the time series.

Proposition 4: _In a rank-order by size the first city grows disproportionately to its size and thus pulls away steadily from its neighbors in the rank-order to become over two to three times the size of the second city and more than three to four times that of the third city._ The test of this proposition will follow Mark Jefferson's convention of assigning an index number of 100 to the size of the "primate" city and of indexing the other two cities accordingly. These relative values will then be examined to determine whether they fall within the ranges postulated above and whether "primacy" does in fact increase over time. In addition the identity of the first city will be ascertained to see if the same city occupies the top rank continuously.

Proposition 5: <u>Taking all cities with populations over 100,000 as a size-class, the rank-order of cities by size exhibits a high degree of stability; this degree of stability is independent of both increases in the absolute number of cities and the rate of growth of all cities.</u> A corollary of this is that there is no significant inverse association between the rank-order of cities by size and their rank order by decade rates of growth. In order to test the original proposition we make use of the Spearman rank-order coefficient of correlation, $= \frac{6\ S\ d^2}{n(n^2-1)}$, where *d* is the difference between a city's rank at some initial point in time and its rank at some later point in time and *n* is the total number of cities in the rank-order; this coefficient ranges in value from -1.00 to +1.00. In this case, where we rank the largest city 1 at both points in time, perfect stability would be indicated by a rho of +1.00. The "city" and "class" methods will be combined in order to consider in parallel a set of time series (of varying lengths): in the case of the U.S. 1890-1900, 1890-1910.........; 1900-1910, 1900-1920...; 1910-1920, 1910-1930........and so on, to 1940-1950. Each country will have such a set of rank-order coefficients computed with the first six census dates of the time series as their initial points. This approach allows us to control for the number of cities by holding them constant, then permitting them to vary and comparing the two outcomes. This will be accomplished by computing the Pearsonian product-moment correlation between the two variables. The same coefficients will be made for the mean rate of growth of all cities vis-a-vis the rank-order coefficients.

Although irrelevant to the validity of the proposition, frequency distributions of rank shifts by magnitude and direction will be presented in order to document rank stability in detail.

To measure the lack of association between population size and decade rate of growth, the rank-order by size at one point in time is compared with the rank-order by growth rate for the decade beginning at that initial point, e.g. an 1890 rank-order by size will be compared with an 1890-1900 rank-order by decade rate of growth, and so on. The proposition is upheld if the rank-order coefficients thus obtained are not significantly different from 0.00 and negative in value. The statistical test of significance in this case involves the use of the Student "t" distribution and the transformation equation, $t = \frac{\sqrt{n} \cdot r}{\sqrt{(1-r^2)}}$, where r is the rank-order coefficient, n is the number of degrees of freedom (here, N, the number of paired comparisons minus 2). Because the rank-orders are an abstraction from the absolute city sizes and rates of growth, the two relationships--size vs. size and size vs. growth rate--will be graphed to make sure that they are linear and non-linear respectively.

Proposition 6: Taking all cities with populations over 100,000 as a size-class, their mean rate of growth tends to retard with time. The test of this propostion will be to compute the mean rates of growth for the first two and the last two decades of the time series, to ascertain the mean values for each of these two decade periods and then to compare these two mean values with each other. Both the "city" and "class" methods will be employed in order to study the effect of increases in the number of cities on the mean rates of growth.

-Findings-

The number of cities in a country varies with the ratio of urban to total population by a constant factor, so that at any one point in time, one may use the value of either quantity to predict the other. In this study the number of cities has been used to predict the ratio of urban total population. Note that the errors are comparable to those for predictions of the number of cities from a knowledge of the ratio; the axes are reversed, and the two parameters -- the slope and y-intercept -- assume different values. In both cases these parameters are empirically derived constants, and the equation is known as Stewart's rule.

The standard errors of estimate are extremely low in India (.8%) and in the U.S. (.9%), not quite so low in Japan (4.4%). This means that if the urban :total ratio is predicted from the number of cities by the rule, 68% of the time these predictions will be accurate within

Table 3

Stewart's rule: the number of cities in relation to the ratio of urban:total population[20]

Country	Year	No. Cities (C)	Urban/Total Pop. (U) observed	computed	Standard Error of Estimate
U.S.A.	1890	1,348	35.1	35.9	
	1900	1,737	39.7	40.8	
	1910	2,262	45.7	46.5	
	1920	2,722	51.2	51.1	
	1930	3,165	56.2	55.0	
	1940	3,464	56.5	57.6	
	1950	4,054	62.7	62.3	
					0.9
Japan	1888	n.a.	n.a.	n.a	
	1898	1,314	32.8	32.6	
	1908	2,027	45.7	47.9	
	1920	2,196	51.6	51.5	
	1930	2,559	60.1	59.0	
	1940	2,616	67.5	60.1	
	1950	3,745	75.2	82.8	
					4.4
India	1891	1,528	n.a.	n.a.	
	1901	2,251	10.0	09.0	
	1911	2,355	09.4	09.3	
	1921	2,675	10.2	10.3	
	1931	3,335	11.1	12.4	
	1941	3,715	12.8	13.5	
	1951	4,542	16.8	16.0	
					0.8

n.a. = not available

±.8% in India, ±.9% in the U.S. and ±4.4% in Japan. The relative smallness of this potential error brings out a salient feature of urbanization, namely that as a higher proportion of a country's total population comes to live in cities, the number of cities correspondingly increases to accommodate them.

Increases in the average size of all cities are positively correlated with increases in the number of cities. Table 4 classifies the absolute figures for these two variables and indicates their correlation. The relationship is clearly evident in the U.S. experience where the product-moment coefficient *r* is +.98. Consequently, 96% ($100r^2$) of the variability in the size of the average city is ascribable to its relation to the absolute number of cities. The coefficients for Japan and India are not so high, +.75 and +.73 respectively. As a result only 56% and 53% of the variation in average city size can be explained by reference to the number of cities. Thus Tisdale's preconditions for urbanization--increase in points of concentration and growth in size of individual points of concentration--are only partially satisfied.

Had it not been for extensive bombing of cities during World War II, the coefficient for Japan would have been higher; as it was, the bombing killed many urban inhabitants and triggered a vast urban-to-rural migration. By 1950 the return to the Japanese cities had not yet compensated for the wartime population losses. In the case of India, the correlation is markedly higher for the last two decades than for the first three when the average city size actually declined. This turn

Table 4

The number of cities with populations over 5,000 in relation to the average size of all such cities

Country	Date	Number of Cities over 5,000 (C)	Average Size of Cities over 5,000 (S)
U.S.A.	1890	1,348	16,399
	1900	1,737	17,418
	1919	2,262	18,633
	1920	2,722	19,942
	1930	3,165	21,876
	1940	3,464	21,556
	1950	4,054	23,404
Japan	1888	n.a.	n.a.
	1898	1,314	10,924
	1908	2,027	11,180
	1920	2,196	13,015
	1930	2,559	15,000
	1940	2,616	18,717
	1950	3,745	16,706
India	1891	1,528	n.a.
	1901	2,251	10,462
	1911	2,355	9,938
	1921	2,675	9,460
	1931	3,335	9,169
	1941	3,715	10,777
	1951	4,542	13,201

	Pearsonian correlation coefficients (S,C)	
U.S.A.	$r = +.98$	$100r^2 = .96$
Japan	$r = +.75$	$100r^2 = .56$
India	$r = +.73$	$100r^2 = .53$

n.a. = not available

of events coincides with the approach to, and beginning of, "take off" in the Indian economy according to Rostow's chronology.

Lorenz curves emphasize the relative distribution of urban population over the range of cities. To index the degree of concentration, forty-five degree tangents are drawn coincident with the Lorenz curve, and the x-intercept of this tangent is interpreted as an index of urban concentration--the higher the index number, the greater the degree of concentration. The values of these x-intercepts are given in Table 5 below.

Table 5

The Index of Concentration: Urban Population

U.S.A.	Japan	India
1890: 55.5	1888: n.a.	1891: 35.3
1900: 55.7	1898: 37.1	1901: 37.5
1910: 61.8	1908: 33.0	1911: 44.0
1920: 62.7	1920: 39.3	1921: 44.5
1930: 63.0	1930: 44.9	1931: 45.1
1940: 61.5	1940: 51.8	1941: n.a.
1950: 61.2	1950: 46.0	1951: 50.3.

n.a. = not available

These indices disclose that urban concentration peaks in the U.S. (1930) and in Japan (1940) whereas it is still on the rise in India. The Japan series is exceptional in two respects. First, the index number declines between 1898 and 1908 as the first, large scale rural-to-urban migrations proceeded not to the largest cities but to the smaller ones. This process is detailed in Table 6.

Table 6

Frequency Distribution:
Number of Cities by Size-Class, 1898-1908

	Number of cities whose population is		
	10,000-50,000	50,000-100,000	100,000 plus
1898	212	13	8
1903	245	16	9
1908	344	19	10.

A combination of circumstances is responsible for this rapid increase in the number of smaller cities and their population. These small cities still had extensive handicraft industries, and the largest cities had not yet become so highly industrialized as to constitute a magnet to rural labor.

Secondly, the 1940-1950 decline in Japan is a reflection once again of the distortion of the data due to wartime damage to Japanese cities. On the other hand, in the U.S. the processes of "suburbanization" and "decentralization" have resulted in a diminution of urban concentration since 1930, a trend that probably will continue for some decades. It is likely that the 1960 index for Japan will disclose a continuation of the trend towards greater concentration. India looks to be undergoing an acceleration of its concentration although the data for 1941 are missing and prevent us from being certain. If the degree of concentration is positively correlated with the timing of economic growth as it appears to be, then the index for Japan may be expected to peak within two decades; this would be in keeping with its thirty-year lag behind the U.S. in

Table 7

THE RANK SIZE RULE

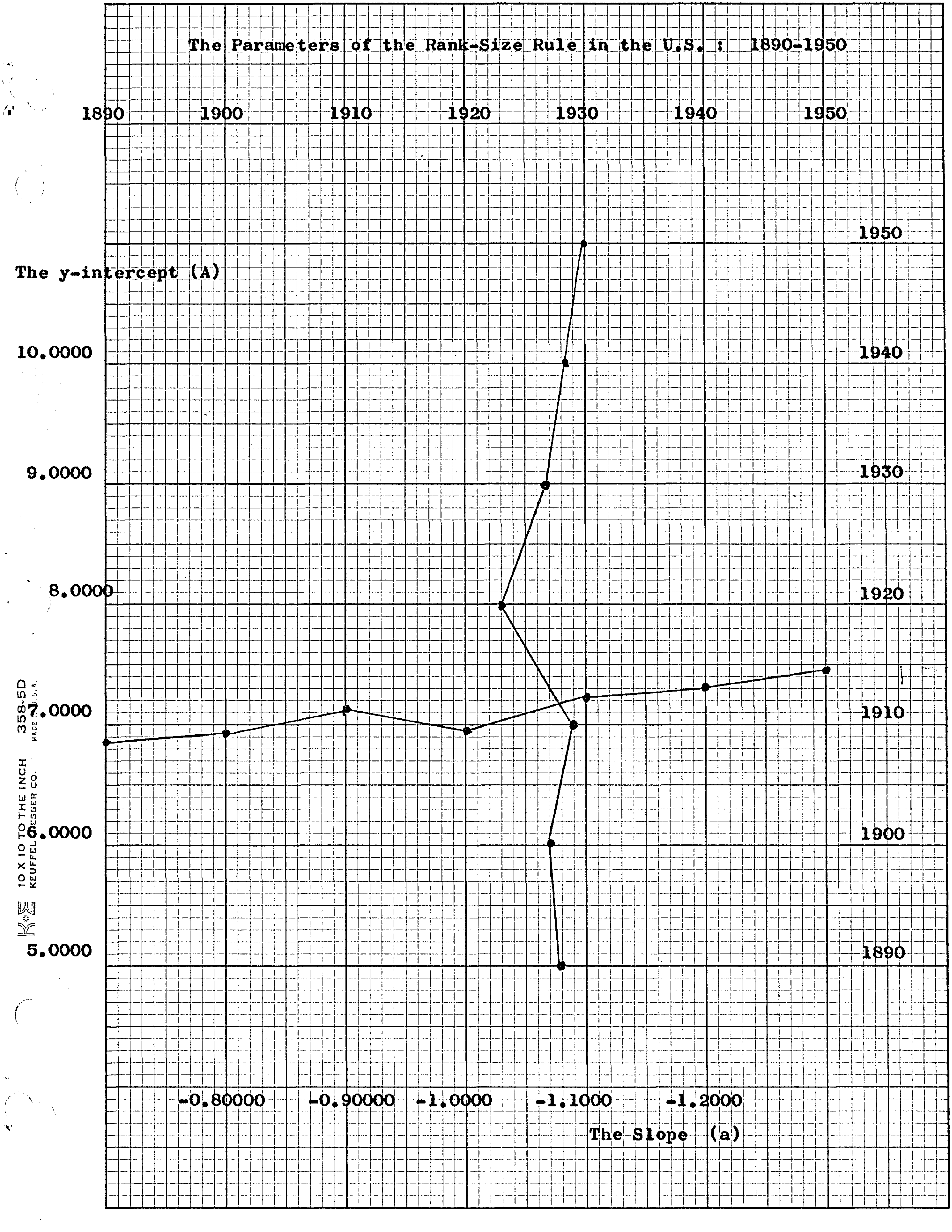
Country	Date	The Average Error in Predicting the Number of Cities by Size-Class
U.S.A.	1890	1%
	1900	2%
	1910	4%
	1920	5%
	1930	13%
	1940	15%
	1950	20%
Japan	1888	15%
	1898	25%
	1908	27%
	1920	24%
	1930	22%
	1940	24%
	1950	25%
India	1891	13%
	1901	16%
	1911	18%
	1921	18%
	1931	26%
	1941	6%
	1951	8%

Table 8

THE RANK-SIZE RULE

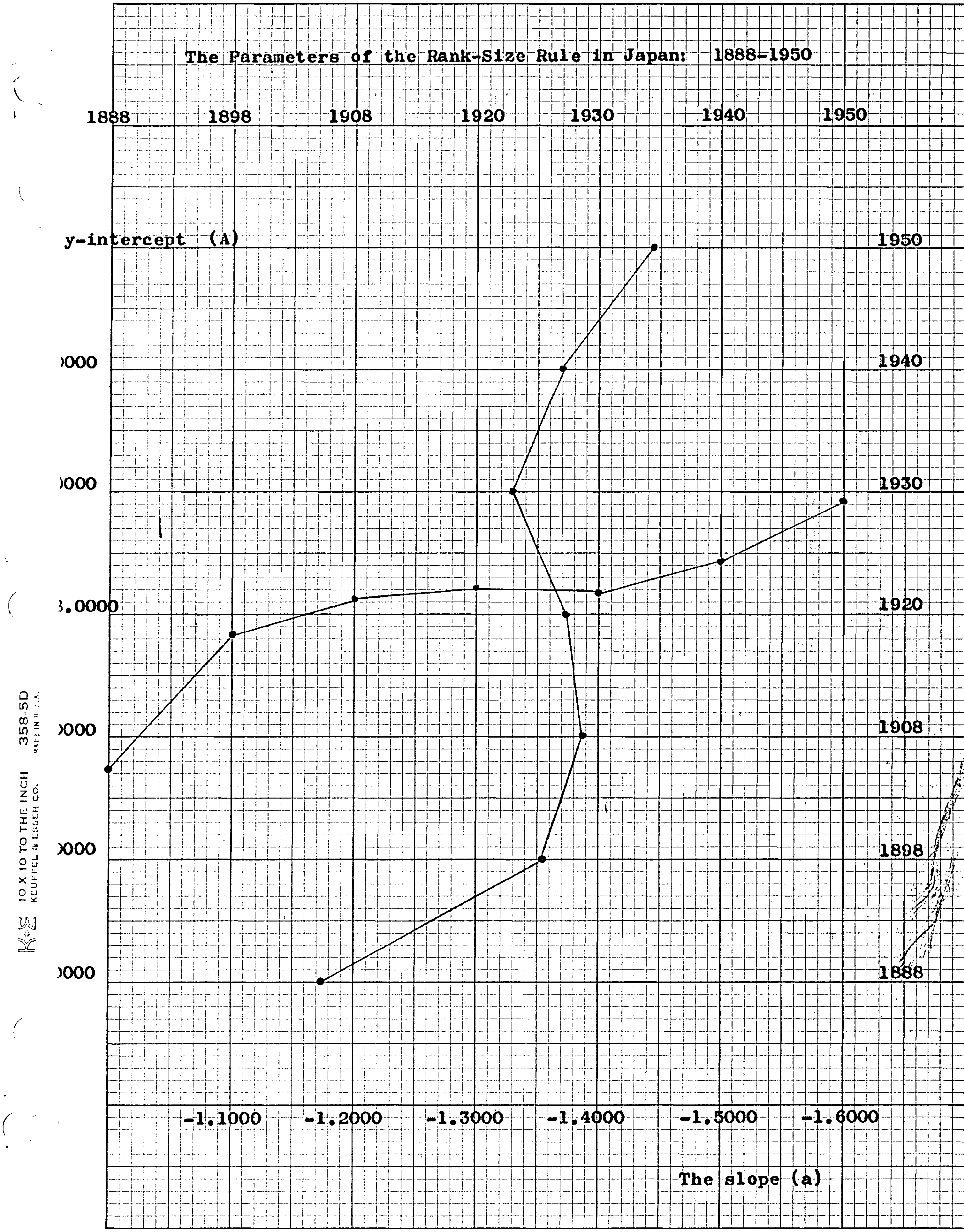
Country	Date	The "M" Constant	The Size of the Top City	% Error
U.S.A.	1890	2,227,000	2,507,400	-11.2%
	1900	2,887,000	3,437,200	-8.4%
	1910	3,563,000	4,766,900	-7.5%
	1920	5,808,000	5,620,000	3.3%
	1930	6,037,000	6,930,000	-12.9%
	1940	5,906,000	7,455,000	-20.8%
	1950	6,386,000	7,892,000	-19.1%
Japan	1888	515,100	1,313,300	-60.8%
	1898	586,300	1,440,100	-59.2%
	1908	701,700	1,626,600	-56.9%
	1920	984,500	2,173,200	-54.7%
	1930	1,412,000	2,070,900	-31,9%
	1940	1,395,000	6,778,800	-79.5%
	1950	1,473,000	5,385,100	-72.7%
India	1891	683,000	773,000	-11.6%
	1901	896,800	776,000	15.5%
	1911	992,700	979,400	1.3%
	1921	1,141,000	1,175,900	-2.9%
	1931	1,270,000	1,161,400	9.3%
	1941	1,700,000	1,695,200	0.0%
	1951	3,041,000	2,839,300	7.1%

Chart 1



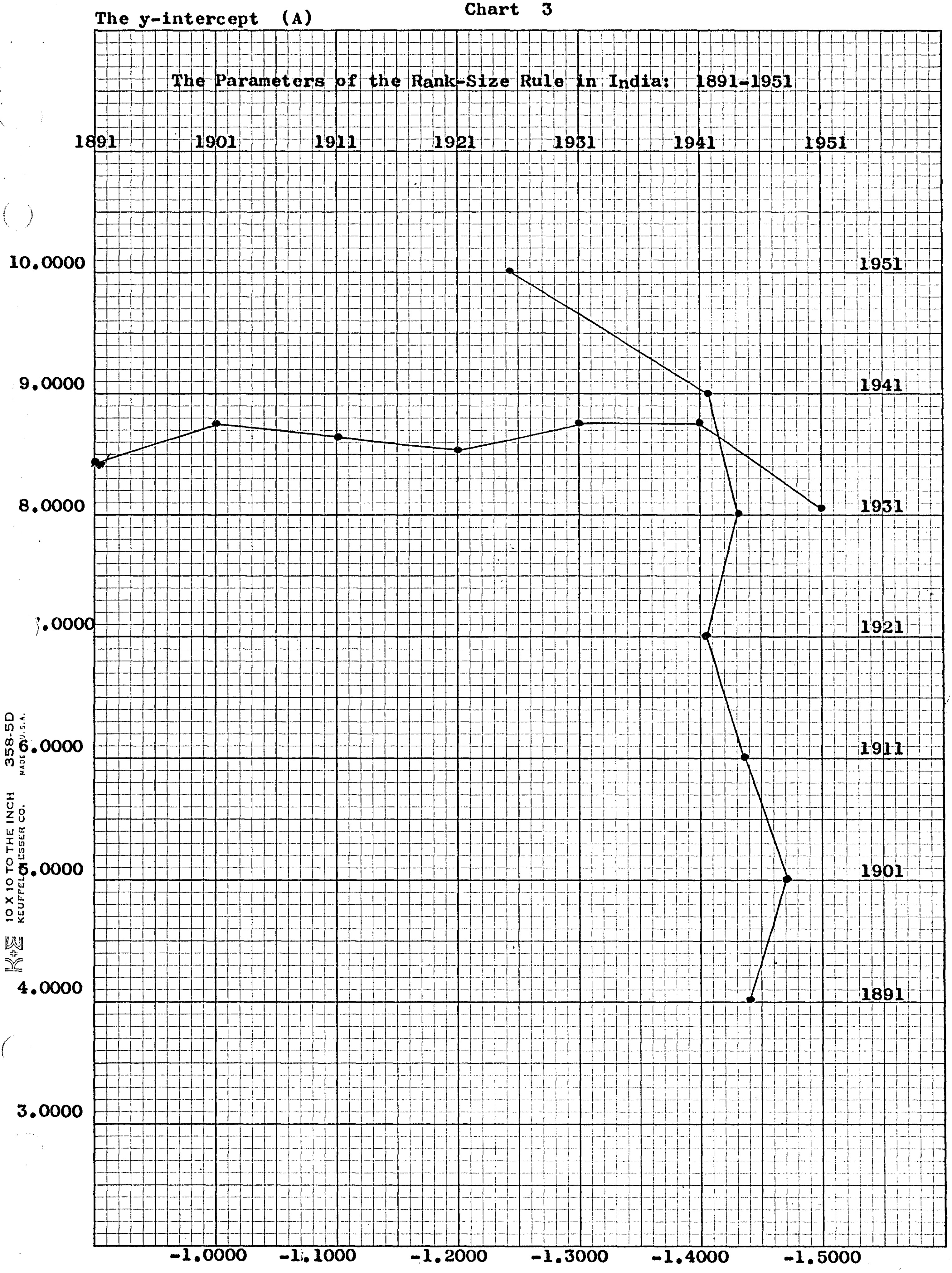
The Parameters of the Rank-Size Rule in the U.S. : 1890-1950
1890
1900
1910
1920
1930
1940
1950
The y-intercept (A)
10.0000
9.0000
8.0000
7.0000
6.0000
5.0000
1950
1940
1930
1920
1910
1900
1890
-0.80000
-0.90000
-1.0000
-1.1000
-1.2000
The Slope (a)
10 X 10 TO THE INCH 358-5D
KEUFFEL & ESSER CO. MADE IN U.S.A.

Chart 2



The Parameters of the Rank-Size Rule in Japan: 1888-1950
1888
1898
1908
1920
1930
1940
1950
y-intercept (A)
)0000
)0000
3.0000
)0000
)0000
)0000
1950
1940
1930
1920
1908
1898
1888
-1.1000
-1.2000
-1.3000
-1.4000
-1.5000
-1.6000
The slope (a)
10 X 10 TO THE INCH 358-5D
KEUFFEL & ESSER CO.

Chart 3


The y-intercept (A)
The Parameters of the Rank-Size Rule in India: 1891-1951
1891
1901
1911
1921
1931
1941
1951
10.0000
9.0000
8.0000
.0000
6.0000
5.0000
4.0000
3.0000
1951
1941
1931
1921
1911
1901
1891
-1.0000
-1.1000
-1.2000
-1.3000
-1.4000
-1.5000
10 X 10 TO THE INCH 358-5D
KEUFFEL ESSER CO. MADE U.S.A.

arriving at "take off" and "technological maturity." India, by this reckoning, is still far from its peak.

The generalized rank-size rule in its logarithmic form, log y=A plus a . log x, provides a good prediction of the distribution of cities by size. The predictions prove not be particularly accurate except for the first half of the U.S. times series (1890-1920) and the last two decades in India (1940, 1950). Further, the U.S. "average error" increases steadily with time which is just the opposite of what we might expect in view of the .98 correlation between number of cities and average size of cities. The Indian figures do exhibit the "expected" relationship in that as the association of those two variables goes up, the accuracy of the rank-size rule becomes markedly greater. The Japanese data are as puzzling as those for the U.S.; they show a consistently high "average error", about 25%, throughout. This evidence refutes the notion that the rank-size rule invariably becomes a better predictor as a country becomes increasingly urbanized, whether that urbanization is defined in terms of number of cities, urban total population ratio, average size of all cities or all three indices together. In the case of the U.S., the findings run directly counter to those of Vining[21] and Madden[22] whose graphic interpretations of the rank-size rule give one the impression of a remarkably stable system. It must be remembered, however, that a plot of city-size distributions on log-log paper suppresses a substantial amount of the variability because of the scale factor between logarithms and absolute values. Perhaps this is what accounts for the discrepancy between the two findings. If so, the

Table 9

THE RANK-SIZE RULE

Dates	The Pareto Coefficients (a) of U.S.A.	Japan	India
1888		1.18	
1890	1.08		
1891			1.44
1898		1.36	
1900	1.07		
1901			1.47
1908		1.39	
1910	1.09		
1911			1.44
1920	1.03	1.59	
1921			1.41
1930	1.07	1.33	
1931			1.43
1940	1.08	1.37	
1941			1.41
1950	1.10	1.44	
1951			1.24

conclusions of Vining and Madden as to the stability of the city size distribution are thrown into question.

If the errors in prediction are examined by size-class, we see that the rank-size rule both overstates and understates the number of "expected" cities at each end of the rank-order. This casts doubt upon Stewart's claim that an S-shaped curve is a better predictor of city size distributions; for that to be the case, the rank-size rule would have to overstate the number of highest-ranking cities throughout and to understate the number of lowest-ranking cities or vice-versa. Stewart's claim cannot be disallowed, however, unless two comparable sets of predictions are contrasted with the data and show the S-curve predictions to be much less accurate.

If Stewart's case were confined to the largest city only, it would perhaps have more validity. When the modified M constants of the generalized rank-size rule are compared with the sizes of the largest cities (see Table 8), they generally understate the absolute value except in India where a long history of colonialism and the existence of two "primate" cities complicate the matter. It may be argued that where "primacy" is shared, the two "primate" are smaller in size than a single "primate" city would be.

The variability of the two parameters of the rank-size rule, A and a, takes place within a relatively narrow range (see Charts 1-3), and in each case this variation may be roughly described by a straight line. The a's are the Pareto coefficients; their variability, although small, is significant in absolute value as well as in the secular trend that they establish. These Pareto coefficients index the relative proportions of large-, medium- and small-sized cities--the smaller the coefficient, the greater the proportion of large cities. Allen made much of the fact that these coefficients fluctuate little in "modern" industrial countries in contrast to the widely held notion that such countries are experiencing growth primarily in their largest cities. The U.S. data bears out his contention (see Table 9). So do the figures for India which indicate a closer approximation to unity within the last two decades, precisely the experience of the U.S. coefficients during the latter half of the 19th century.

The Japanese data are not so unambiguous on the matter. This writer's data conflict with those of Allen. Inasmuch as Allen's procedure was replicated, the reason for the discrepancy is not entirely clear although one possibility is the use of a different number of size-classes in fitting a least-square line to the data. There is agreement on the 1920 coefficient (-1.59), but not on the 1950 figure (-1.44 versus -1.16 according to Allen). If Allen's figure is accepted, then the interpretation of the Indian experience is also applicable to Japan, namely that its coefficients are beginning to approximate unity more closely as the country industrializes more extensively.

In a rank-order by size, the first city grows disproportionately to its size and thus pulls away steadily from its nearest neighbors in the rank-order to become over two to three times the size of the second

Table 10

Primacy: the relative sizes of the three largest cities (size of the No. 1 city = 100)

Country	Date	No. 1 City	No. 2 City	No. 3 City	Name of No. 1 city
U.S.A.	1890	100	44	42	New York
	1900	100	49	38	New York
	1910	100	46	32	New York
	1920	100	48	32	New York
	1930	100	49	28	New York
	1940	100	46	26	New York
	1950	100	46	26	New York
Japan	1888	100	34	21	Tokyo
	1898	100	57	25	Tokyo
	1908	100	75	27	Tokyo
	1920	100	58	28	Tokyo
	1930	100	84	37	Osaka
	1940	100	48	20	Tokyo
	1950	100	36	20	Tokyo
India	1891	100	99	53	Bombay
	1901	100	76	50	Calcutta
	1911	100	94	50	Calcutta
	1921	100	96	45	Bombay
	1931	100	92	51	Calcutta
	1941	100	80	37	Calcutta
	1951	100	90	50	Bombay

city and more than three to four times that of the third city. Nowhere in our data do we find unmistakable evidence of the dynamic that is supposedly at work i.e. disproportionate growth leading to "primacy." According to the proposition (termed the "law of the primate city" by Mark Jefferson), one might expect to see New York out-distance its nearest neighbors in the rank-order, but this is only partially the case. It is only the third city the declines in the "expected" manner from about two-fifths the size of New York to one-quarter its size while the second city maintains its relative position at roughly half the size of New York. The 2:1 ratio that prevails between New York and the second city is also what would be "expected" according to the rank-size rule. The "law of the primate city" and the rank-size rule were formulated independently of each other at about the same time, but this possible degree of overlap between the two of them has not always been recognized or acknowledged.[23]

The data for Japan display a change in direction. First there is a decrease in the first city's "primacy" up to 1930 and subsequently an increase for the last two decades with the net result that the first city returns to its original position vis-a-vis the other two cities. This fluctuation contradicts Mark Jefferson's thesis that "primacy" is essentially an irreversible process. The 1923 earthquake in Tokyo in which 68,000 people died does not appreciably lessen the contradiction because that city's "primacy" was declining even before the earthquake occurred to accelerate the process.

India presents a situation that might be termed "dual primacy" in which Bombay and Calcutta vy with one another for the top rank while the third city is relatively stable at roughly half their size. Mark Jefferson spoke only of deviations from the "law of the primate city". Berry[24] has generalized the concept to allow for "dual primacy," and Ginsberg[25] talks about "regional primacy" which takes us from the nation as a whole to the domain of subnational areas and introduces a spatial dimension into the concept. The idea of "regional primacy" has certain implications bearing on urbanization, but time precludes our investigating this aspect.

With regard to the identity of the first city, only in the U.S. does the first city, New York, maintain its dominant rank throughout. In Japan Osaka overtakes Tokyo between 1920 and 1930 only to be overtaken itself within the next decade; this exchange may perhaps be overlooked inasmuch as the great earthquake of 1923 is most likely responsible for this turn of events. As for India, Bombay and Calcutta interchange positions four times within six decades, but they may be interpreted as "dual primates."

The data tend to support the "law of the primate city" more with regard to relative position than to the dynamics of "primacy." The dynamics would seem to be the more crucial aspect of the two, and if this is the case, the data here cast considerable doubt upon the validity of the "law of the primate city." Finally a change in terminology is clearly called for in that the term "law" denotes invariance which is distinctly not the case.

Taking all cities with populations over 100,000 as a size-class, the rank-order of cities by size exhibits a high degree of stability: this degree of stability is independent of both increases in the absolute number of cities and the rate of growth of all cities. A high degree of stability is found in all three rank-orders by size. The rank-order coefficients are tabulated by time period in Table 11. Out of a total of 63 coefficients, 60 are significantly different from zero at the .01 level. All three exceptions occur in the 1891 time series for Japan: 1891-1911, 1891-1921 and 1891-1931. In absolute terms, 33 of the 63 coefficients are above +.90, 46 above +.85, 54 above +.80 and only 2 below +.75. Tables 12 and 13 document the amount and direction of the few individual rank shifts that do take place. All of this strongly suggests that once a city has grown above 100,000, its relative position in that size-class is unlikely to change very much with time.

This rank-order stability is relatively independent of the mean rate of growth for all cities over 100,000. Table 14 lists the product-moment coefficients between these two variables as +.59 (U.S.), -.70 (Japan) and +.75 (India). If the mean growth rate were to effect rank-order stability, the relationship would be converse i.e. rapid growth would result in diminished rank-order stability. In this context, only the coefficient for Japan is in that "expected" direction, and only about half (49%) of the variation in stability is accounted for by its association with the mean rate of growth. This finding differs from that of Gibbs, but he fails to quantify the correlation and to conclude from other than the extreme cases of his sample, i.e. the highest and lowest

Table 11

Rank-Order by Size Versus Rank-Order by Size

Spearman rank-order coefficients of correlation (ρ)

Country	No. Cities over 100,000	1891-1901	1891-1911	1891-1921	1891-1931	1891-1941	1891-1951
U.S.A.	28	+.95	+.93	+.88	+.87	+.86	+.85
Japan	6	+.89	+.70	+.60	+.83	+1.00	+.94
India	22	+.94	+.84	+.77	+.77	+.77	+.75
			1901-1911	1901-1921	1901-1931	1901-1941	1901-1951
U.S.A.	38		+.93	+.87	+.82	+.81	+.78
Japan	8		+.90	+.90	+.90	+.95	+.98
India	24		+.94	+.87	+.85	+.84	+.80
				1911-1921	1911-1931	1911-1941	1911-1951
U.S.A.	50			+.98	+.96	+.95	+.92
				+.92	+.86	+.91	+.87
				+.93	+.90	+.86	+.87
					1921-1931	1921-1941	1921-1951
U.S.A.	68				+.96	+.94	+.90
Japan	16				+.93	+.80	+.80
India	26				+.96	+.94	+.95
						1931-1941	1931-1951
U.S.A.	98					+.98	+.93
Japan	28					+.78	+.77
India	31					+.97	+.94
							1941-1951
U.S.A.	93						+.97
Japan	45						+.88
India	46						+.96

Table 12

Frequency Distribution of Rank Shifts by Magnitude
(rank-order by size)

Country and Decade	0-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27	28-30	31-33
U.S.A.											
1890-1900	15	8	2	1							
1890-1900	22	9	2	3	1					1	
1910-1920	20	21	7	1	5						
1920-1930	32	14	5	6	7	1					
1930-1940	44	6	18	10	4	1	2	1	1		1
1940-1950	45	23	8	10	3	3	1				
India											
1891-1901	12	6	2	1							
1901-1911	11	8	4	1							
1911-1921	17	4	1	1							
1921-1931	21	4									
1931-1941	24	3	2								
1941-1951	33	7	6								
Japan											
1888-1898	6	0									
1898-1908	8	0									
1908-1920	10	0									
1920-1930	13	4									
1930-1940	19	2	5	1		1					
1940-1950	18	8	5	5	1						

Table 13

Frequency Distribution of Rank Shifts by Direction
(rank-order by size)

Country	Decade	Direction of Rank Shift (by size) (+)	(-)	(0)	Total Number of Cities
U.S.A.	1890-1900	8	15	3	26
	1900-1910	12	21	5	38
	1910-1920	21	22	7	50
	1920-1930	25	37	6	68
	1930-1940	27	53	8	98
	1940-1950	29	56	8	93
Japan	1888-1898	1	1	4	6
	1898-1908	1	1	6	8
	1908-1920	2	3	5	10
	1920-1930	8	5	3	16
	1930-1940	14	12	2	28
	1940-1950	19	19	7	45
India	1891-1901	7	10	4	21*
	1901-1911	10	10	3	23*
	1911-1921	10	11	2	23*
	1921-1931	9	6	10	25*
	1931-1941	10	10	8	28*
	1941-1951	22	17	8	47*

*Data is missing for Trichonopoly

Table 14

Cities over 100,000 Population

Country	Decade	Rank-order correlation: city size vs. city size (ρ)	Population: mean decade rate of growth (x)	Number of Cities: decade rate of growth (C)
U.S.A.	1890-1900	.95	41.6%	34.6%
	1900-1910	.93	45.0%	42.9%
	1910-1920	.98	41.3%	36.0%
	1920-1930	.96	27.2%	29.4%
	1930-1940	.98	3.1%	0.0%
	1940-1950	.97	17.3%	17.0%
Japan	1888-1898	.89	44.3%	33.3%
	1898-1908	.90	42.3%	25.0%
	1908-1920	.92	35.7%	60.0%
	1920-1930	.93	63.3%	75.0%
	1930-1940	.78	93.0%	60.7%
	1940-1950	.88	18.7%	42.2%
India	1891-1901	.94	20.7%	9.1%
	1901-1911	.94	5.8%	0.0%
	1911-1921	.93	11.4%	8.3%
	1921-1931	.96	24.1%	19.2%
	1931-1941	.97	65.2%	56.7%
	1941-1951	.96	69.9%	63.8%

Pearsonian Product-Moment Correlations (r)

	$r(\rho,\bar{x})$	$100r^2$	$r(\rho,C)$	$100r^2$
U.S.A.	+.59	.35	-.45	.20
Japan	-.70	.49	+.68	.46
India	+.75	.56	+.60	.36

rank-order coefficients and the highest and the lowest mean rate of growth, to the neglect of the intermediate cases.

An increased number of cities does not seem to have much impact on the degree of rank-order stability either. The product-moment coefficients in this case are -.45 (U.S.), +.68 (Japan) and +.60 (India) which indicates that the number of cities has even less of an association with rank-order stability than does the mean rate of growth. The highest values are positive; therefore the degree of association that does exist tends to be in the direction of greater rank-order stability as the number of cities increases. Therefore, one reason why a large majority of the rank-order coefficients are so stable is that the number of cities has increased steadily from decade to decade in all three countries. Yet this is but one of the factors involved and from the absolute value of the coefficients, a minor one at that.

A precondition of rank-order stability is that the rates of growth of all cities in the rank-order be either positively correlated with, or unrelated to, their population size. If the correlation is negative, the individual rates of growth in the size-class would converge as would their ranks; thus rank-order stability would break down. Conversely, a positive correlation would mean a persistence of the rank-order by size but with ever increasing gaps between the various city sizes.

The correlations of rank-orders by size with those by rate of growth yield a series of coefficients (see Table 15); none of these deviate significantly from zero at the .05 level except for two decades

Table 15

Rank-Order by Size Versus Rank-Order by Rate of Growth

Spearman rank-order coefficients of correlation (ρ)

	1890-1900	1900-1910	1910-1920	1920-1930	1930-1940	1940-1950
U.S.A.	+.07	-.12	+.12	+.17	+.02	+.02
No. cities	28	38	50	68	98	93
Japan	-.66	-.06	+.09	-.19	-.15	-.45
No. cities	6	8	10	16	28	45
India	+.11	+.13	+.02	-.15	+.20	+.31
No. cities	22	24	24	26	31	46

in Japan (1888-1898 and 1940-1950) and one in India (1941-1951). This finding further corroborates the stability proposition above. In the 100,000 plus size-class rate of growth appears largely unrelated to city size, and therefore rank shifts are relatively rare. If the entire city size distribution were studied, one might find a different situation because of the scale factor. The gap in size between the "largest" and "smallest" cities is much greater, and intuitively it would seem that a city of 5,000 can experience a doubling of population much more readily than New York to take the most extreme case or even a city of 100,000. Whether this scale factor is linear and/or predictable is another matter; it may only hold for certain broad size-classes.

Taking all cities with populations over 100,000 as a size-class, their mean rate of growth tends to retard with time. In the U.S. the mean rates of growth clearly retard from the beginning of the time series to its end. The degree of retardation is greater when the number of cities is varied rather than being held constant. This variation is documented in Table 14 which shows that the U.S. mean rate declines about 33% in the forty years with increasing numbers of cities whereas it drops only about 22% if the number of cities is held constant. The Indian cities do not experience retardation; in fact, their mean rate of growth increases substantially over the forty-year time span. Further, this increase is greater (54%+ versus 33%+) when the number of cities is allowed to increase--precisely the opposite of what took place in the U.S. time series. If it had not been for the depressed growth rates in the decade encompassing

World War II, the Japanese cities would have had a similar experience. As it is, the results for Japan are mixed with the "city" method registering a decline and the "class" method an increase in mean rate. (While the absolute number of cities seems to have a divided effect on the mean rate of population growth, the rate of increase in the number of cities has a more positive association with retardation. That is, a high growth rate in the number of cities correlates with a high rate of growth for the population, and conversely low rates with low rates--see Table 16. This lends support to Madden's thesis that cities compete with one another for population within a closed system and that increased competition usually results in a general lowering of the growth rates.)

Neither the Japanese nor the Indian cities have reached their peak in terms of growth, but arrival at a turning point and subsequent retardation seem to be an inevitable prospect. Boulding[26] notes that retardation is a general characteristic of growth no matter what the particular phenomenon at hand. If anything in nature were to grow indefinitely at an increasing rate, it would soon engulf all of the world within its bounds. Such a state of being has been envisioned by Doxiadis with his concept of ecumenopolis and by Fuller with his idea of a world-world town and his map of the world as virtually one, continuous land mass. But this is urbanization in a radically different sense than the traditional definitions that have governed this study.

Table 16

Type of method employed		Decade rates of growth: all cities over 100,000					
		U.S.A.					
		1890-1900	1900-1910	Mean	1930-1940	1940-1950	Mean
Population	"city"	27.4%	30.0%	28.7%	3.4%	9.0%	6.2%
	"class"	41.6%	45.0%	43.3%	3.1%	17.3%	10.2%
No. Cities		34.6%	42.9%	38.7%	0.0%	17.00%	8.5%
		Japan					
		1888-1898	1898-1908	Mean	1930-1940	1940-1950	Mean
Population	"city"	54.9%	51.1%	53.0%	71.3%	17.6%	44.5%
	"class"	44.3%	42.3%	43.3%	93.0%	18.7%	55.8%
No. Cities		33.3%	25.0%	29.2%	60.7%	42.2%	51.3%
		India					
		1891-1901	1901-1911	Mean	1931-1941	1941-1951	Mean
Population	"city"	15.5%	2.9%	9.2%	43.0%	41.9%	42.5%
	"class"	20.7%	5.8%	13.3%	65.2%	69.9%	67.6%
No. Cities		9.1%	0.0%	4.5%	56.7%	63.8%	59.7%

-Conclusions-

While the process of urbanization is too complex to be defined in terms of a single variable, its character may be adequately described by a few key variables and their permutations. These key variables include total population, number of cities, city sizes, and city rates of growth. In most cases two variables are correlated with each other by means of parameters empirically derived from the data (propositions 1 and 3), of a product-moment correlation (proposition 2) or of a rank-order correlation (proposition 5). There are two cases where a direct comparison is made at two different points in time between index numbers (proposition 4) and mean values (proposition 6).

Stewart's rule (proposition 1) relates number of cities to ratio of urban population (sum of city sizes) to total population. Tisdale's preconditions (proposition 2) correlate number of cities with average city size or urban population divided by number of cities. The rank-size rule (proposition 3) links city sizes with the rank-order by city size. The "law of the primate city" (proposition 4) compares city size with city size and city rate of growth with city rate of growth at the very top of the rank-order by size. The notion of rank-order stability (proposition 5) contrasts rank-order by size with rank-order by size at two different points in time; the measure of stability itself is correlated with mean rates of growth for both urban population and number of cities. Finally, retardation (proposition 6) compares the levels of the mean rates of growth of urban population at the beginning and end of the time series which are in turn compared with the rates of increase in the number of cities.

These six propositions then all derive from the four key variables above, and they constitute a series of indices by means of which one can measure the degree of urbanization in any given country. Our findings have made clear that a particular country does not necessarily register an advance along all of these fronts simultaneously besides which a few indices of urbanization experience a change in direction in their secular trend. For example, in India, the average size city actually declined at the same time that both number of cities and urban population were on the increase. In the U.S., the index of urban concentration and mean rates of population growth peak and then decline, yet Stewart's rule continues to hold, Pareto coefficients remain near unity and rank-order stability persists at a high level.

As a result any overall measure of urbanization has to be more on the order of a series of net resultants (of various index values) than of a series of "stages" narrowly defined. The process of urbanization has a time-dimension by definition, and it is the change in the variables over time that enables us to describe urbanization. The "stage" theorists, however, such as McKenzie[27], are as guilty of oversimplification as those who define urbanization along a single dimension; the indices that describe urbanization simply do not fit into a neat set of time categories.

An urban analogue to the "stage" theory of economic development is a tempting prospect, but one encounters difficulty in such reasoning by analogy. First, the validity of the economic "stage" theory has itself come into question following critiques by North[28] and Perloff[29]. Secondly, numerous sections of the economic "stage" theory do not have their urban counterparts; instead of self-sustained growth, there is retardation and discontinuous rates of growth from decade to decade, to take one example. Where an analogy might be drawn is from Perloff's work which documents a variety of possible growth sequences.

The above does not imply that one cannot compare the performances of various countries in time, but it does suggest that any attempt to combine various indices into a "stage" theory of urbanization faces serious, if not insurmountable, obstacles. A review of the several indices of this study brings out certain time lags with regard to Japan and India; the indices differ in their absolute value and sometimes in their direction over time. The U.S. ranks as the most highly urbanized country of the three on virtually all of the indices. India lags from between forty to an indeterminate number of years behind the U.S. If the 1940-1950 distortions are discounted, Japan is intermediate between the two for the most part. These comments have to be qualified with the statement that the time series spans only sixty years, and some secular trends may be obscured in the particular time-slice we have studied. Nonetheless, the data does suggest certain sequences. Retardation has occurred in the U.S., the mean rate of growth has slowed down in Japan and there is no immediate sign of a slowdown in India. Similarly, the Lorenz curve index peaks first in the U.S., appears about to peak in

Japan and still shows no sign of reaching a turning point in India. The correlations between number and size of cities are skewed with highest value in the U.S. and lowest in India: further the decline in average city size occurred much earlier in Japan than in India and was of shorter duration. By the time our time series starts the Pareto coefficient for the U.S. had already dropped close to unity and stabilized; the coefficient for Japan began its decline in 1920 while the one for India starts to decline between 1940 and 1950.

But not all the indices disclose such phased time lags. The rank-size rule results are mixed, there is loss of predictability in the U.S., loss and then gain in predictability in India and relatively no change in the high "average error" in Japan. In rank-order stability by size, Japan's cities demonstrate the least stability and many more negative correlations between city size and city rate of growth even though the absolute values are not great. It may be that India will find itself in a comparable position once industrialization gains momentum there.

On a comparative basis the propositions present the following picture. Stewart's rule (proposition 1) finds confirmation in all three countries. Tisdale's preconditions for urbanization (proposition 2) manifest themselves continuously in the U.S. series and probably also in date for Japan if the wartime distortions are discounted but not in India except the last decades. While the rank-size rule (proposition 3) does not, for the most part, accurately describe city size distributions, considerable rank-order stability (proposition 5) prevails among the individual cities in all three countries. Jefferson's "law of the primate city" (proposition 4) finds only partial confirmation, and then with some qualification. It provides a reasonably good prediction of relative position in the U.S. and in India (with "dual primacy" as a qualification), but there is almost no evidence of disproportionate rates of growth in the data. Only the first and third cities in the U.S. grow steadily apart. Finally, Madden's concept of retardation is confined to the U.S. experience; rates of growth have yet to peak in Japan and India although the Japanese rate shows some sign of nearing its turning point.

Having examined the findings in their temporal and comparative aspects, it is time to place this study in its larger context--within the conceptual framework of a "system of cities" as used by a number of writers[31]. The concept of a "system of cities" implies, at a minimum, that there is a certain order underlying the group of cities in a country. There have been suggestions from some quarters that this order extends to the cities of the world taken as a whole. The nature of this underlying order has been subject to several interpretations in the literature. One school of thought, exemplified in the writings of Rutledge Vining[32], confines itself to an analysis of the statistical behaviour of cities in the aggregate. Cities as a group exhibit certain mathematical properties which make a prediction of their performance characteristics possible; frequently, these predictions are couched in probabilistic terms.

Another school of thought subscribes to the "nodal-linkage" model and takes as its central concern the functional interdependency of individual cities within the system. What cities "do" i.e. their functions has an impact upon what cities "are" i.e. their demographic, characteristics, etc. The empirical studies of Duncan[33], Harris[34] and Ullman[35] among others are representative of this school. These works have their theoretical counterpart in the writings of "general systems theory" which place much stress on the fact that changes in one part of a closed system lead to changes in all other parts of the system. While interdependency demands a closed system, it is a fact that "systems of cities" are becoming more and more open both at the national level as the number of autarkic regional economies diminishes and at the international level as international trade barriers fall, common markets expand, and supranational political bodies come into being. This latter set of factors is making it increasingly imperative to study urbanization on a worldwide basis.

In addition to differences in the level of explanation (statistical interpretations of aggregate behavior versus functional interpretations of individual city behaviour), there is a difference in the way individual cities are ordered in the system. Vining rank-orders cities along a continuum by size whereas Christaller[36], Brush[37] and others classify cities into discrete, functional size-classes which they regard as somehow "real" or "true" divisions. Vining does not rank-order cities along more than one dimension e.g. size or income, but the "functionalists" do in two separate contexts when they speak of a) the "functional prerequisites" of cities by size-class or type and b) the "functional correlates" of city size. Because there is more than one dimension involved, only a "weak" ordering is possible. Furthermore, "levels, ranks or classes in a hierarchic scheme should be regarded either as conceptual fictions suitable for manipulation in abstract discourse or as categories of convenience for handling empirical data."[38] The principle is more important than the criterion which rests upon it.

In addition to a vertical ordering, the "system of cities" has a horizontal dimension, that is extent in space. The spatial distribution has been interpreted statistically by Vining, Madden and others. Chrostaller, Losch, and Stolper among others have developed functional interpretations of the same phenomenon. Most of these writings have homogeneity as their premise but no one has yet to control systematically for homogeneity versus heterogeneity in resource endowment and/or socio-economic characteristics.

This study falls largely into the statistical-behavioural category. We have empirically determined certain demographic relationships without recourse either to their functional prerequisites or to the functional correlates of city size. Having ascertained the demographic outcomes of the process of urbanization, we are in a position to work "backward" in the sense of analyzing the dynamic processes that have led up to the outcomes, or "forward" in the sense of linking the process of urbanization to the process of economic development. Berry's study[39] made a beginning in this direction but with oversimplified accounts of both urbanization and industrialization; it will be necessary to detail the process of economic development to a greater extent than Berry has done and certainly to include more than the rank-size rule and the "law of the primate city" as descriptors of the process of urbanization before a satisfactory theory encompassing both processes can be devised. Future research will hopefully elucidate and elaborate upon their association; in this connection, the outcomes that this study establishes might be used as benchmarks in the determination of interdependency.

The limitations of this study will perhaps point to areas which might profitably be researched in the future. First, the 5,000 person minimum in this study may not accurately reflect the critical values in the three countries. The impact of the lower size limit of cities needs detailed study; one useful experiment would be to test Stewart's rule with 10,000, 15,000, 20,000, 25,000, 50,000, 75,000 and 100,000 as alternative lower limits in order to discover the impact these minimums have on the generalization.

The performance characteristics of cities need to be studied systematically by size-class; class limits should be varied to discover the most significant cutting points with regard to various dimensions such as rate of growth. It is not sufficient to limit observations to the 100,000 plus size-class because certain relationships hold there that may not apply to other size-classes; in addition, if the entire spectrum of cities is observed, we may come up with findings opposite to those for the 100,000 plus size-class or alternatively more pronounced.

The use of legal-political cities has introduced a bias into the study whereby the existence of agglomerations of individual cities is overlooked. The propositions may or may not apply with greater force to such agglomerations; Gibbs' conclusion that legal-political cities are just as useful as metropolitan areas in comparative work needs testing in a context similar to this study. This is difficult to do inasmuch as comparable data for metropolitan areas[40] only exist for two dates, 1950 and 1960, and many of the propositions both derive from, and need to be tested by, time series data.

This leads us into the next area of possible improvement--the expansion of the time series. Sixty years is a perhaps not quite long enough a period from which to generalize with much certainty. Several empirical relationships have been based upon six points, a small enough number to allow for considerable error in the fitting of a least-square line or in the computation of correlation coefficients. The particular time slice chosen (1888-1950) may very well have biassed our results; it certainly has precluded the controlling of certain variables while allowing others to vary.

Perhaps most importantly, the three countries studied are not a representative sample of the total universe. Because conditions differ so much from one part of the world to another, it may be impossible to select a truly representative sample and necessary instead to study the universe of countries as a whole. Certainly, the widespread rural-to-urban migrations taking place on all continents make such comprehensive study highly desirable and increase the importance of separating out the universal from the particular. Every study to date has had to resolve the conflicting demands for breadth (number of countries) versus depth (number of cities) versus length (number of years or decades). Resources have rarely been allocated among these three demands in a satisfactory way. Computer technology now makes it possible to program a worldwide study of urbanization that would include all of the cities in each country and a time series for as long as reasonably accurate data are available; this would seem to be the next logical step in the study of urbanization.

The intense nature of urbanization in this century should not obscure the fact that the pace of urbanization was radically different in the past, that historically the process of urbanization has been neither inevitable nor irreversible and that urbanization is but one of many processes that coexist in a country. Nonetheless, as countries becoming increasingly urban, greater knowledge of the performance characteristics of their "system of cities" becomes fundamentally important to an understanding of those societies. This study was formulated as a first step in that direction.

FOOTNOTES

1. Vining, D.R., "Observation and Description in an Analysis of an Economic Problem: The Case of Underdevelopment as an Example," unpublished manuscript, 1962, pp. 121-122.

2. Gibbs, J.P., Urban Research Methods (Princeton: D. Van Nostrand Company, Inc., 1961). Chapters 11 and 12, p. 397.

3. Tisdale, H., "The Process of Urbanization", Social Forces, 20, March 1942, p. 311-316.

4. Vining, R. <u>op</u>. <u>cit</u>. p. 61.

5. Singer, H.W., "The 'Courbes des Populations' A Parallel to Pareto's Law", Economic Journal, 46, June 1936, p. 259.

6. Ibid, p. 258

7. Allen, G.R., "The 'Courbes des Populations': A Further Analysis", Bulletin of the Oxford University Institute of Statistics, 16, May-June 1954, p.184.

8. Ibid p. 186-187.

9. Stewart, C.T., Jr., "The Size and Spacing of Cities", Geographical Review, 48, April 1958, p. 226.

10. Jefferson, M., "The Law of the Primate City", Geographical Review, 29, April 1939, pp. 228.

11. Stewart, C.T., Jr.,<u>op</u>.<u>cit</u>., p. 231

12. Ibid, p. 84.

13. Gibbs, op. cit., p. 454.

14. Ibid, p. 451.

15. Madden, C.H., The Growth of Cities in the United States: An Aspect of the Development of an Economic System, unpublished Ph.D. thesis, U. of Virginia, p. 39.

16. Vining, R., <u>op</u>. <u>cit</u>., p. 64-65.

17. Rostow, W.W., The Stages of Economic Growth Cambridge: Cambridge U. Press, 1960, p. 38.

18. Ibid, p. 59.

19. Gibbs, <u>op</u>. <u>cit</u>., p. 427, italics mine.

Footnotes (continued)

20. The data for Tables 3-14 have been obtained from the following five sources:

    Beale, E.G. Jr., "Population of the Cities of Japan," translation from the Japanese in the Far Eastern Quarterly, 3, August 1944.

    Historical Statistics of the U.S.: 1790-1950 (General Printing Office, 1957).

    Ishii, R., Population Pressure and Economic Life in Japan (Chicago: U. of Chicago Press, 1937).

    Great Britain India Office, Statistical Abstracts Relating to British India, 1891-1941.

    India Central Statistical Organization, Statistical Abstract, India: 1951.

21. Vining, R., "A Description of Certain Spatial Aspects of an Economic System", Economic Development and Cultural Change, 3, 1955, pp. 147-195.

22. Madden, C.H., "On Some Indications of Stability in the Growth of Cities in the United States", Economic Development and Cultural Change, IV, 3, April 1956, pp. 236-252.

23. Jefferson, M., "The Law of the Primate City", Geographical Review, 29, April 1939, and G.K. Zipf, National Unity and Disunity: The Nation as a Bio-Social Organism (Bloomington: The Principia Press, Inc., 1941).

24. Berry, B.J.L., "City Size Distribution and Economic Development" in Economic Development and Cultural Change, 9:4, Part I, July 1961, p. 578.

25. Ginsberg, N., Atlas of Economic Development (Chicago: U. of Chicago Press, 1961), p. 36-37.

26. Boulding, K., "Toward a General Theory of Growth", in J.J. Spengler and O.D. Duncan, Population Theory and Policy (Glencoe: The Free Press, 1956).

27. McKenzie, R., The Metropolitan Community (New York: McGraw-Hill, 1933).

Footnotes (continued)

28. North, D.C., "Location Theory and Regional Economic Growth," Journal of Political Economy, 63, pp. 243-258.

29. Perloff, H., Regions, Resources and Economic Development (Baltimore: The Johns Hopkins Press, 1960).

30. Ibid.

31. Vining, op. cit.; Hoover, E.M., "The Concept of a System of Cities: A Comment on Rutledge Vining's Paper," Economic Development and Cultural Change, 3, January 1955, pp. 196-198; Berry, B. and A. Pred, Central Place Studies: A Bibliography of Theory and Applications (Philadelphia: Regional Science Research Institute, 1961); and Duncan, O.D., et. al, Metropolis and Region (Baltimore: The Johns Hopkins Press, 1960), Chapter 3.

32. Vining, op. cit.

33. Duncan, op. cit.

34. Harris, C.,,"The Market as a Factor in the Localization of Industry," Annals of the American Association of Geographers, 44, 1954, pp. 215-348.

35. Ullman, E.A., American Commodity Flow (Seattle, U. of Washington Press 1957).

36. Christaller, W., Die zentrallen Orte in Suddendeutschland translated by C.W. Baskin, "A Critique and Translation of W. Christaller's Die zonetralen Orte in Suddendeutschland," unpublished Ph.D thesis, U. of Virginia, 1957.

37. Brush, J.E. "The Hierarchy of Central Places in Southwestern Wisconsin," Geographical Review, 43, 1953, pp. 380-402.

38. Duncan, op. cit., p. 51.

39. Berry, op. cit.

40. Davis, K., The World's Metropolitan Areas (Berkeley: U. of California Press, 1959).

Bibliography

A. Books

Abstract of the 13th Census of the U.S., 1910 (Washington: General Printing Office, 1912).

Abstract of the 14th Census of the U.S., 1920 (Washington: General Printing Office, 1923).

Abstract of the 15th Census of the U.S., 1930 (Washington: General Printing Office, 1933).

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

Christaller, W., Die zentralen Orte in Suddendeutschland, translated by C.W. Baskin, "A Critique and Translation of W. Christaller's Die zentralen Orte in Suddendeutschland", unpublished Ph.D. thesis, U. of Virginia, 1957.

Davis, K., The Population of India and Pakistan (Princeton: Princeton University Press, 1951), Chapters 15, 16 and 20.

Duncan, O.D. and A. Reiss, Jr., Social Characteristics of Urban and Rural Communities, 1950 (New York: John Wiley and Company, 1956), Chapter 2.

Gibbs, J.P., Urban Research Methods (Princeton: D. Van Nostrand Company, Inc., 1961), Chapters 11 and 12.

International Urban Research, The World's Metropolitan Areas (Berkeley: U. of California Press, 1959).

Ishii, R., Population Pressure and Economic Life in Japan (Chicago: U. of Chicago Press, 1937).

Madden, C.H., The Growth of Cities in the United States: An Aspect of of the Development of an Economic System, unpublished Ph.D thesis, 1954, U. of Virginia.

Siegel, S., Nonparametric Statistics for the Behavioural Sciences New York: McGraw-Hill Book Company, 1956), Chapter 9.

Hoyt, H., "Is City Growth Controlled by Mathematics or Physical Laws?", Land Economics, 27, August 1951, pp. 259-262.

Jefferson, M., "The Law of the Primate City", Geographical Review, 29, April 1939, pp. 226-232.

Lampard, E.E., "American Historians and the Study of Urbanization", American Historical Review, LXVII, 1, October 1961, pp. 49-61.

Madden, C.H., "On Some Indications of Stability in the Growth of Cities in the United States", Economic Development and Cultural Change, IV, 3, April 1956, pp. 236-252.

Madden, C.H., "Some Spatial Aspects of Urban Growth in the United States", Economic Development and Cultural Change, IV, 4, July 1956, pp. 371-387.

Madden, C.H., "Some Temporal Aspects of the Growth of Cities in the United States", Economic Development and Cultural Change, VI, 2, January 1958, pp. 143-170.

Mandelbrot, B., "Final Note on a Class of Skew Distribution Functions: Analysis and Critique of a Model due to H.A. Simon", Information and Control, Vol. 4, Numbers 2-3, September 1961, pp. 198-216.

Singer, H.W., "The 'Courbes des Populations' A Parallel to Pareto's Law", Economic Journal, 46, June 1936, pp. 254-263.

Stewart, C.T., Jr., "The Size and Spacing of Cities", Geographical Review, 48, April 1958, pp. 222-245.

Stewart, J.Q., "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population" in J.J. Spengler and O.D. Duncan, Demographic Analysis: Selected Readings", pp. 344-371. (Glencoe: The Free Press, 1946).

Taeuber, I., "Urbanization and Population Change in the Development of Modern Japan", Economic Development and Cultural Change, IX, 4, October 1960.

Tisdale, H., "The Process of Urbanization", Social Forces, 20, March 1942, pp. 311-316.

Simon, H., Models for Man -- Social and Rational (New York: John Wiley, 1957), Part II, Chapter 9.

Taeuber, I., The Population of Japan (Princeton: Princeton University Press, 1958), Chapters III, IV, V and VIII.

U.S. Census of Population, Volume I, Characteristics of the Population, Part A, Number of Inhabitants (Washington: Government Printing Office, 1961).

Yule, G. and M. Kendall, An Introduction to the Theory of Statistics (London: Charles Griffin and Company, 1949), Chapters 13 and 17.

Zipf, G.K., National Unity and Disunity: The Nation as a Bio-Social Organism (Bloomington: The Principia Press, Inc., 1941).

B. Journal Articles

Allen, G.R., "The 'Courbes des Populations': A Further Analysis", Bulletin of the Oxford University Institute of Statistics, 16, May-June 1954, pp. 179-189.

Beale, E.G. Jr., "Population of the Cities of Japan", translation from the Japanese in the Far Eastern Quarterly, 3, August 1944, pp. 303-362.

Beckmann, M.J., "City Hierarchies and the Distribution of City Size", Economic Development and Cultural Change, 6, 1958, pp. 243-248.

Berry, B.J.L. and W. Garrison, "Alternate Explanations of Urban Rank-Size Relationships", Annals of the Association of American Geographers, 48, 1958, pp. 83-91.

Boulding, K., "Toward a General Theory of Growth", in J.J. Spengler and O.D. Duncan, Population Theory and Policy (Glencoe: The Free Press, 1956).

Davis, K. and H.H. Golden, "Urbanization and the Development of Pre-Industrial Areas", Economic Development and Cultural Change, 3, October 1954, pp. 6-26.

Hoover, E.M., "The Concept of a System of Cities", Economic Development and Cultural Change, 3, 1955, pp. 196-198.

Vining, R., "Delimitation of Economic Areas: Statistical Conceptions in the Study of the Spatial Structure of an Economic System", Journal of the American Statistical Association, 18, 1953, pp. 44-64.

Vining, R., "A Description of Certain Spatial Aspects of an Economic System", Economic Development and Cultural Change, 3, 1955, pp. 147-195.

Vining, D.R., "Observation and Description in an Analysis of an Economic Problem: The Case of Underdevelopment as an Example", unpublished manuscript, 1962.