The Relation Between Rents and Incomes, and the Distribution of Rental Values

By W. C. HELMLE

SYNOPSIS: Many parts of telephone plant, such as central office buildings and equipment, conduits, underground and aerial cable at the time of installation must have the capacity to handle not only the immediate demand for telephone service, but also to take care of growth for a number of years to come. In order to engineer such items of telephone plant economically it is necessary to know in advance as accurately as possible what the demand for telephone service will be five, ten, or twenty years in the future. Forecasts of the future market are very necessary for plant engineering, operating plans, rate treatment, and other purposes, in multi-central office cities. In such cities detailed estimates are made of the market some twenty years ahead and of its telephone development under stated rate conditions. Such estimates are called *commercial surveys*, and they involve a study of the various factors which, in the course of events, will be likely to control the industrial, commercial and residential development of the city concerned.

In the course of such a survey, a rental classification of all families is obtained and at the same time a record is made of existing telephone service in each rental class. The rent data of this article have been gathered in representative large cities throughout the country and the results as here set forth are being used together with many other kinds of data to guide the engineering of future additions to the plant of the Bell Telephone System.

In general the income of a family is an index of the market it creates for various commodities including telephone service. Rental values may also be considered as such an index and the present study seeks to correlate rents with incomes. Rents can be readily recorded and classified, whereas it is not feasible to determine the money incomes of large numbers of families. While it may be ideally possible, by a study of rent data, to compare the inherent markets for telephone service and also the strength of the telephone habit in various cities, there are many practical limitations to such a procedure. Comparison of the residence market for telephone service in different cities, as determined by rent values, is made difficult by the fact that the variation between cities in rentals paid for substantially similar dwellings is considerably greater than the variation in prices for food or clothing. Further, there is considerable variation in rent levels even in different sections of any one city. Attempts to compare rent distributions by application of the usual statistical measures of dispersion and skewness have proved unsatisfactory. However, a method of charting has been found by which rent distributions may be readily compared with one another and an index of spread or dispersion determined. It has been found that cumulative curves of rent distribution may be plotted on logarithmic probability paper to yield straight lines for a large number of cities. These are called logarithmic skew distributions. Although it has not been found possible to assign any special significance to the particular value of the index of rent dispersion in any city, this index appears to remain practically constant for that city regardless of changes in the level of prices. In the appendix the mathematical features of the logarithmic skew curve are discussed.—*Editor.*

IT is a well recognized fact that the better class families, *i.e.*, those with higher incomes, are a better market for telephone service than the poorer families. For purposes of market analysis in commercial surveys it is not feasible to determine the money incomes received by families but the rental values of dwellings, which, as will be shown, are a measure of the incomes of their occupants, are comparatively readily collected and classified. Rent data obtained in the course of a commercial survey show the "character" of a city and are used as a basis for estimates of the future residence telephone market.

In view of the importance of these rent data, it seems desirable to study them in some detail to find out just what their limitations are.

There may be set down in advance certain things which it is desirable to know, as, for instance, the relationship between money incomes and house rents and methods and limitations of comparison of different cities on the basis of rental values. On the first point, as applied to any particular city, a knowledge of the relation between incomes and rents is desirable in a general way, although there is no necessity to translate the telephone market expressed in terms of rent types into a scale of incomes. On the second point, the comparison of different cities, it should be ideally possible by a study of rent data to compare the inherent economic markets for telephone service, and also to measure differences in the strength of the telephone habit, but in practice only rough approximations may be made.

Certain limitations to work of this kind are fairly evident. The most obvious difficulty is the fact that rent levels have changed along with the general price level. Rent levels in various cities differ according to the varying degrees of housing congestion and the varying social standards of the population. Furthermore, the variation in rent levels extends to different sections of any one city. The mere fact that a given family paid say $30 rent is not an indication of that family's economic condition or its value as a telephone prospect, unless there is also known the city and the part of the city in which that family lived, and the time when the given rent was paid. Therefore rent data from different cities and of various dates are not directly comparable at their face value. To adjust the money values of house rents for an accurate comparison of the telephone markets in different cities would require a knowledge of the relative proportions of income spent for rent in the different cities, of the relative levels of incomes and rents at the time of the surveys as compared with their levels in some base year, and perhaps of other factors equally difficult to estimate.

Various rent tabulations can not be compared one with another without knowing something of the way in which rents are distributed about their average. The nature of the distribution is determined by the house count data, but from those data in their usual form it is not easy, when making comparisons, to make proper allowances for differences in rent levels and in the schedule of rent classes. In the following pages there is discussed a method of charting by which rent distributions may be readily compared, and their spread or dispersion determined.

The Relation Between Rents and Incomes

Rents as a Market Index. The relation between rents and incomes is concerned with the use of rental values both as an index of telephone market in a given city, and in comparing the markets in different cities. In what follows it is not always possible to separate these two views, but the distinction should be borne in mind by the reader.

The goal of an analysis of residence telephone market is to determine the future sales possibilities. In theory either incomes or rents may be considered as an index of the telephone market. The market index adopted in commercial surveys is the rental of dwellings. This may be considered either as a direct measure of the ability and desire of families to subscribe to telephone service, or as an indirect index, if incomes are considered the real measure of the market. If the first viewpoint is accepted, it may be logically concluded, although not proved, that rents are a better index of telephone market than are incomes. Incomes, as measured in money, are the nearest approach which may be made to a measure of the position of families on an imaginary scale of economic welfare. An attempt to translate rent data to an income basis, as a working method in commercial surveys would introduce errors with no compensating advantage, but a translation of this kind is more or less unconsciously made in making comparisons.

Sources of Information. Much of the literature on the question of house rents versus incomes is generalization based on limited or antiquated data. Such careless statements as "rent approximates about one-third of the average worker's income," may be found in the literature of the subject. Adam Smith, the father of Political Economy, "made the assertion, surprising to us in these days, that the proportion of income spent in house rent is highest among the rich." Frederick Engels concluded in 1857 that rent was 12 per cent and heat and light 5 per cent of the workingman's expenditure, regardless of the amount of his income.

Investigation of budgets in recent years has been confined almost entirely to the field of the wage earning class. The first really comprehensive study was made by the United States Bureau of Labor

Statistics in several states in 1901 and 1902, and is detailed in the 1903 annual report of that organization. It consisted chiefly of a study of 11,156 so-called "normal" families, each including a husband at work, a wife, not more than 5 children all under 14 years and no lodgers or servants. The average income of these families was $651. Original work on a smaller scale has been done by R. C. Chapin (1908) and by the Philadelphia Bureau of Municipal Research (1918). The Bureau of Labor Statistics collected a large amount of data in 1918 and 1919[1] concerning the incomes and expenses of 12,837 families in 92 towns having an average income of $1491. This investigation included families of wage earners and low salaried men, but none of the slum or recent immigrant classes. Families of the lowest type are automatically excluded from such studies as this by their inability to supply the desired information from accounts or from intelligent estimates.

Distribution of Family Expenses. Representative distributions of family expenditures are given in Table I. The National Industrial Conference Board has adopted for use in computing their cost of living index and representative budgets a list of standard weights made by combining the results of a number of studies made from 1901 to 1917. Most importance was assigned to the first Bureau of Labor Statistics study, the results of which it closely resembles. The standard weights used by the Bureau of Labor Statisics are the result of surveys made in 22 cities from July 31 to November 30, 1918, covering families whose average income was $1,434.

TABLE I.

Distribution of Total Family Expenditure

	Averages Found in		Standard Weights Used by	
	Bur. Lab. Stat. First Study 1901–1902	Bur. Lab. Stat. Second Study 1918–1919	Bur. Lab. Stat. Weights in Cost of Living Index	Nat'l Ind. Conf. Board Budgets and Index of Cost of Living
Food	43.13%	38.5%	38.2%	43.13%
Rent	18.12	13.3	13.4	17.65
Clothing	12.95	16.5	16.6	13.21
Fuel and Light	5.69	5.3	5.3	5.63
Sundries	20.11	26.4	26.4	20.38

From Table I it may be inferred that the per cent spent for rent is reduced in a period of inflated prices, at least during the first part

[1] *Monthly Labor Review*, May–December, incl., 1919.

of that period. This is reasonable, since rents respond less rapidly than most other prices to fluctuations in the general price level. An extreme example of this type is found in Germany where rents, which are to some extent under government regulation, "at the present time absorb not more than 3½ per cent of total expenditure as against 20 per cent before the war." [2]

The percentage distribution of total expenses depends on the size of the family, the income received and the city lived in. Of course, it must be understood that any particular family may differ widely from general averages. Other things being equal, large families spend more for food and clothing and less for rent and sundries than do small families. Large families of the lower middle class accommodate themselves to whatever housing accommodations they can afford after the more inflexible demands for other things have been provided for. Less than one room per person is considered over-crowding and the recent Bureau of Labor Statistics investigation found this condition to exist rarely, except in families having more than three children. Families with one to three children were found to have 1.0 to 1.3 rooms per person in almost all cities.

Amount of Income vs. Per Cent Spent for Rent. The extent to which the distribution of expenses is modified by the amount of income received is known only within the very limited range for which data are available. The best recent figures are those of the 1918-1919 study of the Bureau of Labor Statistics. These are given here for 12,096 white families in 92 cities and towns:

TABLE II.

Income	Per Cent of Total Expenses Spent for					
	Food	Clothing	Rent	Fuel and Light	Furniture and Furnishings	Misc.
Under $900	44.1	13.2	14.5	6.8	3.6	17.8
$900–$1200	42.4	14.5	13.9	6.0	4.4	18.7
$1200–$1500	39.6	15.9	13.8	5.6	4.8	20.2
$1500–$1800	37.2	16.7	13.5	5.2	5.5	21.8
$1800–$2100	35.7	17.5	13.2	5.0	5.5	23.0
$2100–$2500	34.6	18.7	12.1	4.5	5.7	24.3
$2500–up	34.9	20.4	10.6	4.1	5.4	24.7

When the original data are examined in detail, it appears that in almost every city as incomes increase the per cent spent for rent

[2] M. Elsas, *Economic Journal*, September, 1921, p. 332.

and food decreases and the per cent for clothing increases. The decrease in the per cent for food as incomes increase is slight and the increase in the per cent for clothing is especially marked in the higher incomes within the range covered. Thus, it appears that among families of moderate incomes as incomes rise the increase is spent by preference for clothing rather than for food or rent. The relative decrease in expenditure for rent as incomes increase is significant in rental analysis. This means that while a 10 per cent difference in rents among the lower rents in a city indicates an average difference in income of about 10 per cent, a similar difference among the higher rents indicates a difference in income of much more than 10 per cent.

Rent Levels in Various Cities. As nearly as may be determined from the Bureau of Labor Statistics data, there is no regular tendency for Eastern, Western or Southern cities to differ from the average of all cities, either in the amount of wage-earners' incomes or in amounts spent for food or clothing. In Southern cities somewhat less is paid for rent than in other cities. This refers only to white families. Negro families have smaller average incomes than white families and at any given income they spend less for rent, more for food and, to a less degree, more for clothing than white families. The size of a city, so far as may be told from these data, does not determine either total incomes or expense for food, rent or clothing.

In different cities the difference in rent levels, that is, variation in rentals paid for substantially similar dwellings, is considerably greater than the difference in levels of prices for food or clothing. The variation in price levels is about twice as great for rents as for food or clothing, reckoned as percentages of the amounts spent for each class. The expenditure for food by the lower middle class families included in this investigation is more nearly the same in different cities than is the expenditure for rent or for clothing. Food expense is the only one of these classes in which all cities are as closely grouped as in total expenses, considering deviations from the averages on a percentage basis. The amounts spent for rent show relatively wide variation between cities. It appears that if a workman moves from one city to another to secure increased wages a large proportion of the increase in income goes for increased rent. This is to be expected since land rents and, to a less extent, construction costs are peculiar to each individual city, much more than food or clothing costs.

A comparison of rent data from the 1918-1919 investigation of the Bureau of Labor Statistics and data from Commercial Surveys leads to the conclusion that differences in average rents in various cities are due at least as much to differences in the level of prices for rents as to differences in the grade of the population. Wage-earners and low salaried people of the types studied by the Bureau of Labor Statistics occupy about the same position in the community in a large number of cities. As a rule they pay about 80 to 90 per cent of the median[3] rent in any city. Exception must be made in the case of cities having an unusually large proportion of negro or very low grade white population. It is interesting in this connection to compare wage rates for different classes of labor in various cities. The variation between cities in wage rates for common labor is proportionately much greater than the variation in wages for work requiring some skill, such as bricklaying and structural iron work.

As examples of the impossibility of accurately rating the grade of a city's population by its median rent alone, we may take four cities where surveys were made in 1921. Spokane and Houston had practically identical median rents of $23.00 and $23.40 respectively, but Houston is not as good a telephone market as Spokane. In Cleveland and Minneapolis the median rents were found to be $35.50 and $31.00 respectively, but this is no measure of the grades of the two cities.

Rent Data from Various Sources, Including England. Some additional rent data is presented here without extended comment. The two following tables show the proportion which rent bears to total expense in different communities.

TABLE III.

Pre-War Expenditures for Rent with a "Normal" Standard of Living

(Senate Report on "Woman and Child Wage Earners")

Manhattan	20.7%
Fall River	17.6%
Georgia and North Carolina	6.3%
Homestead, Pa	15.5%

TABLE IV.

Allowances for Rent in Post-War Standard Workingman's Budgets

	Date	Per Cent for Rent	Rent
No. Hudson Co., N. J.	Jan., 1920	13.5–14.1	$18.00–$19.00
Cincinnati	May, 1920	15.6	22.00
Lawrence	Nov., 1919	13.1–14.1	15.00– 19.50
Fall River	Oct., 1919	9.2–11.6	9.75– 15.20
Philadelphia	Nov., 1919	16.6	.....
United Mine Workers	Dec., 1919	10.2	.....
Washington, D. C.	Aug., 1919	13.3	.....

[3] See Appendix.

The first four of these "standard" budgets are by the National Industrial Conference Board, and the others in order by the Bureau of Municipal Research (Philadelphia), Professor W. F. Ogburn, and the U. S. Bureau of Labor Statistics.

It has already been mentioned that the per cent spent for rent shows a tendency to decrease with increasing incomes. This trend is confined by data from other sources, as follows:

TABLE V

Per Cent of Income Spent for Rent at Different Income Levels

Income	Philadelphia Bur. Mun. Res. 1918 260 Families	Chapin's N. Y. Study 1907–1908 391 Families	U. S. Bur. Lab. Stat. 1901–1902 11,156 Families
$400– $500	......	26.8%	18.6%
500– 600	......	25.9	18.4
600– 700	20.5%	23.6	18.5
700– 800	17.6	21.9	18.3
800– 900	18.1	20.7	17.1
900– 1000	15.8	19.0	17.6
1000– 1100	16.4	18.1	17.5
1100– 1200	14.3	16.2	......
1200– 1300	14.7	19.8	......
1500– 1600	12.3	16.3	......
1900 up	10.2	......	......

Although no data are available for families above the lower middle class, the relationship may be extended by conjecture into the higher income levels. That this is reasonable is brought out in subsequent pages in comparing distribution curves of rental values and incomes.

Some interesting conclusions from English experience are given by Sir J. C. Stamp.[4] The rent corresponding to an income of £160 averages at least £5 greater in London than outside that city. Among the lower incomes, say up to £1000, the variation or dispersion of the percentages paid for rent becomes less as the amount of the income rises. Owner-occupants live in larger houses than tenants with the same total income. It was not found, as is generally supposed, that professional men pay relatively more for rent than business men. The following table is from the work above mentioned;

[4] J. C. Stamp, "British Incomes and Property," 1916.

TABLE VI

Income	*Relative Amount Paid for Rent*
£200–£250	1.0
300– 400	.8
500– 750	.7
1000–1500	.5

Income	*Rent*	*Per Cent Rent*
£160	£28	17.5%
400–£500	40–£50	10
4000	200	5

The second of these two tables represents average conditions for Great Britain.

Rent Distributions

In making comparisons of survey rent data it is desirable to distinguish differences in price levels as they affect rents, differences in economic grade of the population, and differences in the distribution of families about their average grade. Failure to take account of these three factors will result in misleading impressions, which may be illustrated by summaries from successive surveys in Atlanta. The following table shows composites of private residences, flats and apartments:

TABLE VII

Rent Classes	No. of Families		Per Cent Increase
	1913	1920	
$75 up	793	3199	303
55–$75	1129	3582	217
40– 55	2045	4196	105
25– 40	4802	9713	102
20– 25	4197	4725	13
15– 20	5688	4757	—16
10– 15	6881	7223	5
Under $10	21064	15023	—29
Total	46599	52418	12.5

It might be inferred from this set-up that the condition of the poorer families had been very much improved or that the average family had attained a higher condition of well-being. It will be shown later that there was no material change in the distribution of rental values about an average rent when rents are considered as percentages of that average, and it is probable that the principal, if not the sole cause of the changes shown in the table above, is the general rise in the level of prices.

Methods of Study—Graphic Representation. The most convenient and practical method of studying rent distributions is by the use of graphs and charts. The distribution of values of rents or of other variables may be charted in either a detail or a cumulative form. A detail curve shows at any value of the independent variable the frequency of occurrence of items of that value. A cumulative curve shows at any value of the variable the number (or better, the per cent) of all cases which have values below (or above) that value. Cumulative curves are better than detail curves for presenting rent data since the number of classes into which the data are divided is small and the class widths are non-uniform, resulting in uncertain

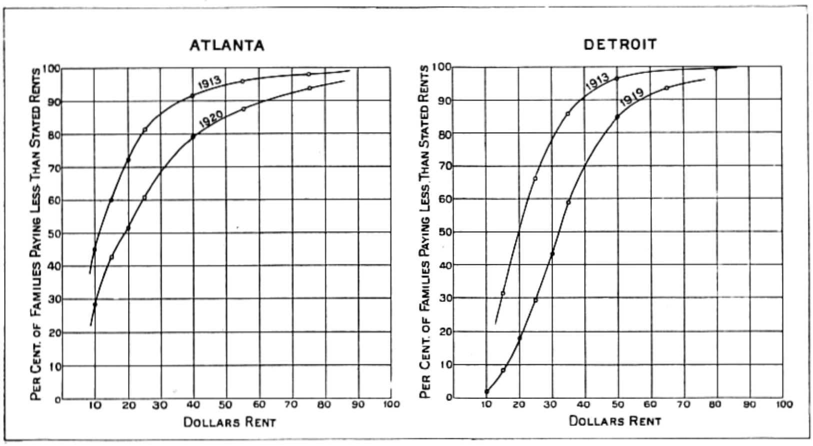

Fig. 1

curves of the detail type.[5] Attempts to compare rent distributions by application of the usual statistical measures of dispersion and skewness have proved unsatisfactory.

Typical rent distributions plotted in the cumulative manner on ordinary coördinate paper are shown in Fig. 1. Diagrams of this type may be used to determine the rent paid by families of corresponding position in the rent scale at the dates of successive surveys, but they do not give a very clear picture of changes in the distribution of rents and from them it is not readily apparent whether rents are closely concentrated or widely distributed in any given case.

[5] Detail curves for rent and income distributions are most easily drawn on paper with a logarithmic scale both ways.

Logarithmic Probability Charts. Cumulative curves for rent distributions may be plotted on logarithmic probability paper[6] in which case the resulting graph is a straight line for a large number of cities. Such a graph will be said to represent a *logarithmic skew distribution.* In the appendix there is given a discussion of frequency curves, with special reference to curves of this type. The essential point in reading charts on logarithmic probability paper is that the slope of the line determines both the spread or dispersion of the data and the skewness or lack of symmetry of distribution. Since the horizontal scale is logarithmic it follows that the dispersion is represented on a percentage and not a linear basis. A steep slope indicates a close concentration of the data, a less steep slope indicates a wider distribution, and parallel lines indicate distributions which are identical on a percentage basis. As explained in the appendix, the most convenient index or coefficient for expressing the spread or dispersion of a distribution is the ratio of the upper quartile[7] to the median rent. If the curves for a given city are closely parallel for successive surveys it follows that there has been no material change in the character of the distribution. In other words, rents have increased approximately proportionately at all points of the scale.

Examples of charts of this kind (Figs. 2-4) are shown for twelve cities for which successive surveys are available. Curves for successive surveys are nearly parallel in eight of the twelve cities. For Cleveland, Dallas and Houston there are distinct differences in the curves for the two dates, indicating changes in the distribution of rents, which changes may be measured since horizontal distances between points on the curves for two dates represent the percentage increases in rents.

When a rent distribution is plotted on logarithmic probability paper the points do not always lie on a straight line, but a straight line of best fit may be chosen by eye, giving greatest weight to points near the middle of the scale of ordinates. Of the rent distributions for large cities which have been plotted on this paper, nearly one-third are very closely represented by straight lines, an equal number are slightly concave upward, and the remainder are more or less concave downward. Most of the deviations from straight lines are slight. The examples submitted herewith (cities in which successive surveys have been made) are rather poorer than the average in this

[6] See an article by G. C. Whipple in the *Journal of the Franklin Institute* for July and August, 1916, for a description of this paper and some examples of its use in the field of sanitation.

[7] See Appendix.

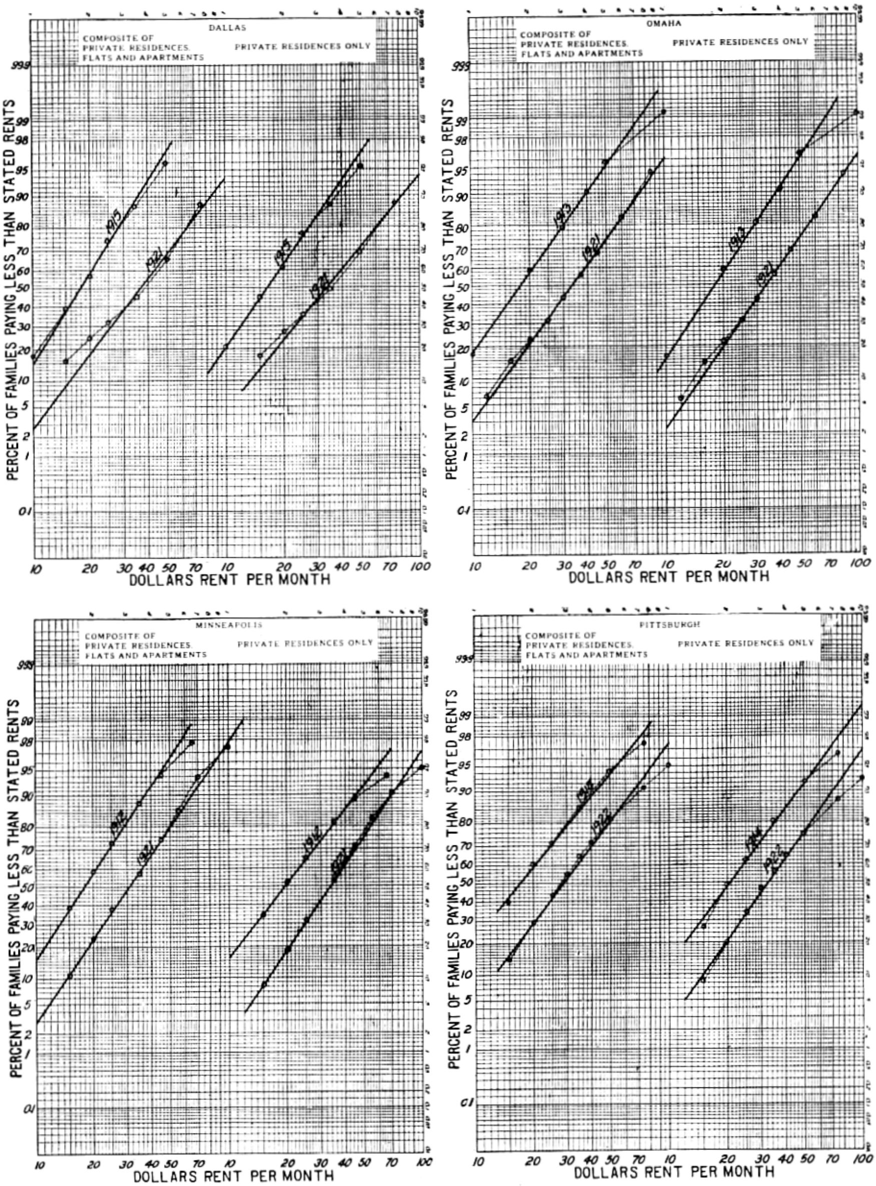
DALLAS
COMPOSITE OF PRIVATE RESIDENCES, FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH
OMAHA
COMPOSITE OF PRIVATE RESIDENCES, FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH
MINNEAPOLIS
COMPOSITE OF PRIVATE RESIDENCES, FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH
PITTSBURGH
COMPOSITE OF PRIVATE RESIDENCES, FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH

Fig. 2

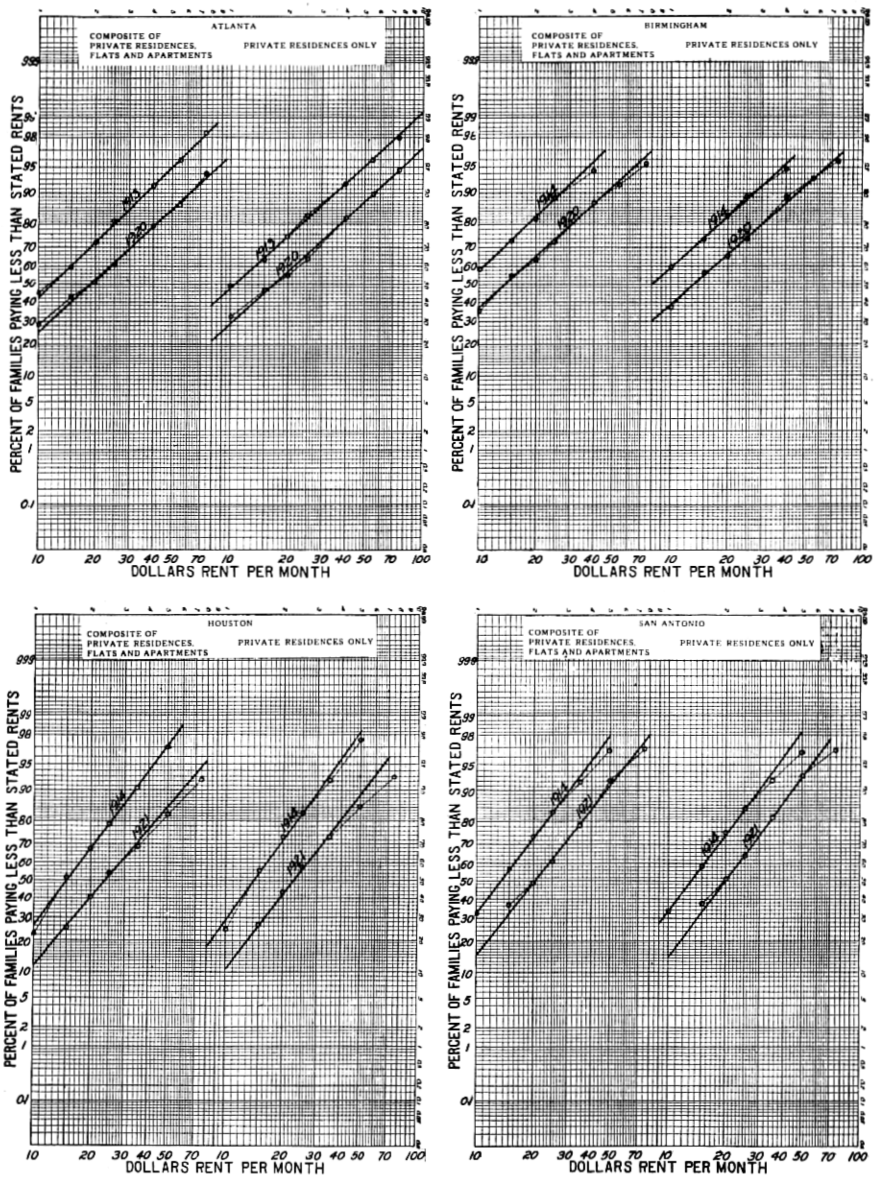
ATLANTA
COMPOSITE OF PRIVATE RESIDENCES, FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH
BIRMINGHAM
COMPOSITE OF PRIVATE RESIDENCES, FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH
HOUSTON
COMPOSITE OF PRIVATE RESIDENCES, FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH
SAN ANTONIO
COMPOSITE OF PRIVATE RESIDENCES, FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH

Fig. 3

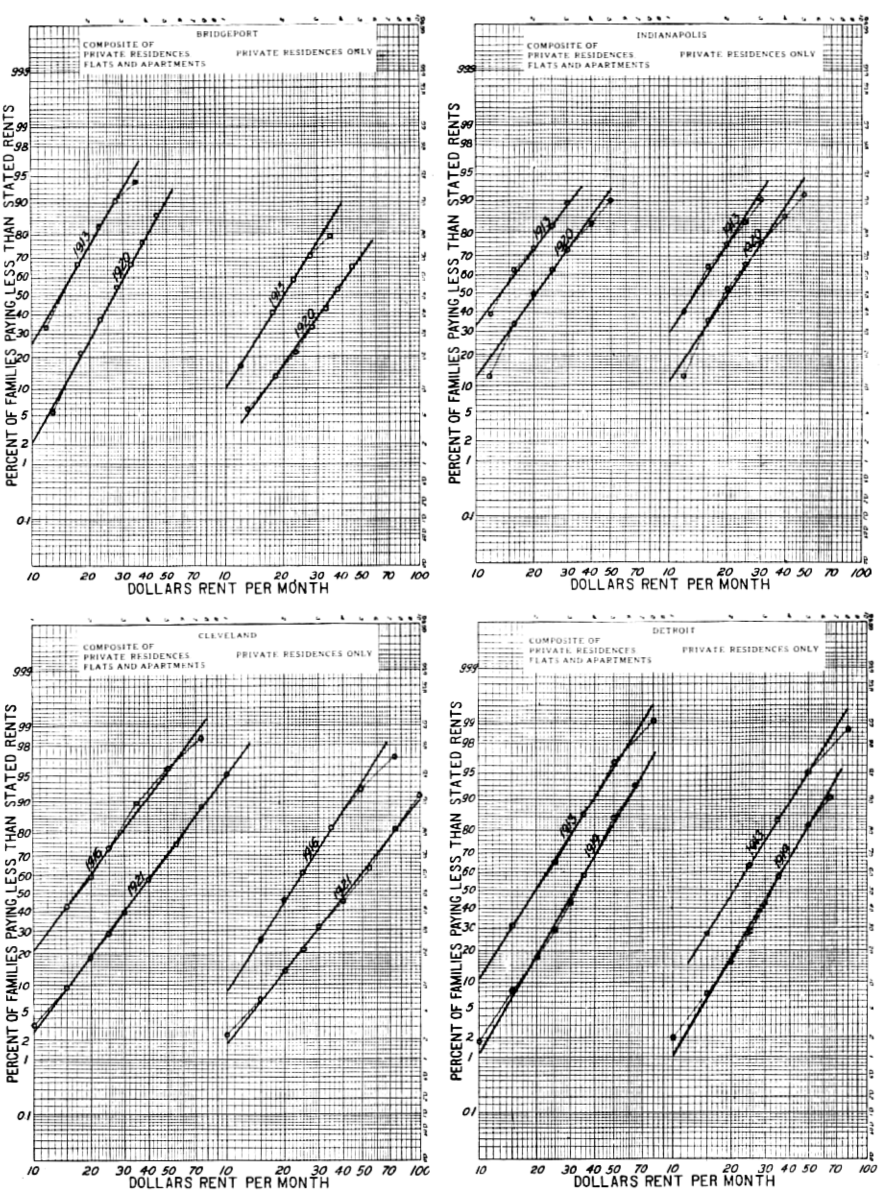
BRIDGEPORT
INDIANAPOLIS
CLEVELAND
DETROIT
COMPOSITE OF
PRIVATE RESIDENCES
FLATS AND APARTMENTS
PRIVATE RESIDENCES ONLY
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
DOLLARS RENT PER MONTH

Fig. 4

respect. Concavity upwards represents a distribution which is less skew than the theoretical logarithmic skew curve, and concavity downwards a distribution of greater skewness. No great importance can be assigned to small differences of this sort, as they are not permanent between successive surveys, whereas the general type of the distribution is quite constant for a given city.

The justification for assuming that rents follow the logarithmic skew curve is made stronger by certain data from Volume 19 of the report of U. S. Immigration Commission made in 1912. This commission collected a large mass of data concerning the living conditions of families of the immigrant type. The data are classified by nationality of head of family, by income, etc. Distributions of amounts paid for house rent per apartment, per room and per person for certain nationalities are shown in Fig. 5. The data shown were chosen from those classes which were made up of the largest numbers, and the deviations from straight lines shown by data for other groups are in both directions, so the straight line relation may be considered fairly representative. It may be noted that the rents per month per person show a greater dispersion than the rents per room or per apartment. These latter moreover show as small a spread as do rents for any of the cities studied as a whole.

Fig. 5 also shows the distribution of British house rents at intervals during the period 1890-1913. There has been a gradual but steady decrease in the dispersion of rents during the period covered. Unless the relation between rents and incomes has radically changed, this means that the inequality of distribution of wealth has been decreased, and that the condition of the poor has been improved as compared to that of the rich. Data for 1830 indicate that the inequality of distribution was distinctly greater at that date than in 1890. Changes in the relative condition of the rich and poor may be readily demonstrated by charts of this kind, but of course conclusions regarding absolute degrees of well-being must be reached by other means.

Distribution of Rents Compared with that of Incomes. Significant conclusions regarding the relation between rents and incomes may be drawn from a comparison of their respective distributions. Fig. 6 shows a detail curve for income distribution in the United States based on preliminary data of the National Bureau of Economic Research. These data are subject to revision but are the best available and are sufficiently accurate for comparative purposes. The usual way to chart income distribution assumes conformity with Pareto's law which says that the frequency curve of incomes may be plotted as a straight line on double logarithmic paper, either on a

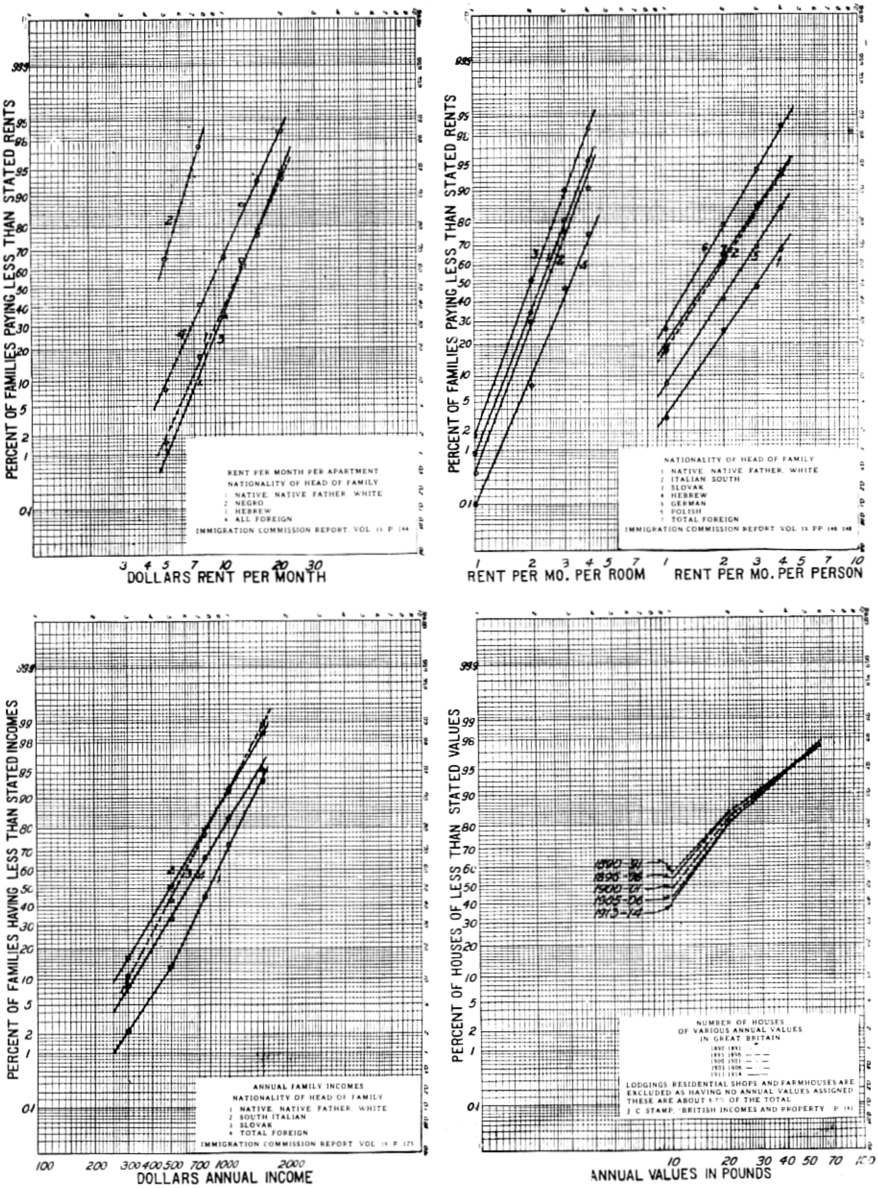
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
RENT PER MONTH PER APARTMENT
NATIONALITY OF HEAD OF FAMILY
1 NATIVE, NATIVE FATHER, WHITE
2 NEGRO
3 HEBREW
4 ALL FOREIGN
IMMIGRATION COMMISSION REPORT, VOL. 15, P. 144
DOLLARS RENT PER MONTH
PERCENT OF FAMILIES PAYING LESS THAN STATED RENTS
NATIONALITY OF HEAD OF FAMILY
1 NATIVE, NATIVE FATHER, WHITE
2 ITALIAN, SOUTH
3 SLOVAK
4 HEBREW
5 GERMAN
6 POLISH
7 TOTAL FOREIGN
RENT PER MO. PER ROOM
RENT PER MO. PER PERSON
PERCENT OF FAMILIES HAVING LESS THAN STATED INCOMES
ANNUAL FAMILY INCOMES
NATIONALITY OF HEAD OF FAMILY
1 NATIVE, NATIVE FATHER, WHITE
2 SOUTH ITALIAN
3 SLOVAK
4 TOTAL FOREIGN
DOLLARS ANNUAL INCOME
PERCENT OF HOUSES OF LESS THAN STATED VALUES
1890-91
1895-96
1900-01
1905-06
1913-14
NUMBER OF HOUSES
OF VARIOUS ANNUAL VALUES
IN GREAT BRITAIN
LODGINGS, RESIDENTIAL SHOPS AND FARMHOUSES ARE
EXCLUDED AS HAVING NO ANNUAL VALUES ASSIGNED
J. C. STAMP, "BRITISH INCOMES AND PROPERTY"
ANNUAL VALUES IN POUNDS

Fig. 5

detail or cumulative basis. This law does not hold for the lower income levels which may be best represented by a curve of approximately hyperbolic form, as shown in Fig. 6. The same income data are shown plotted on logarithmic probability paper in an insert on

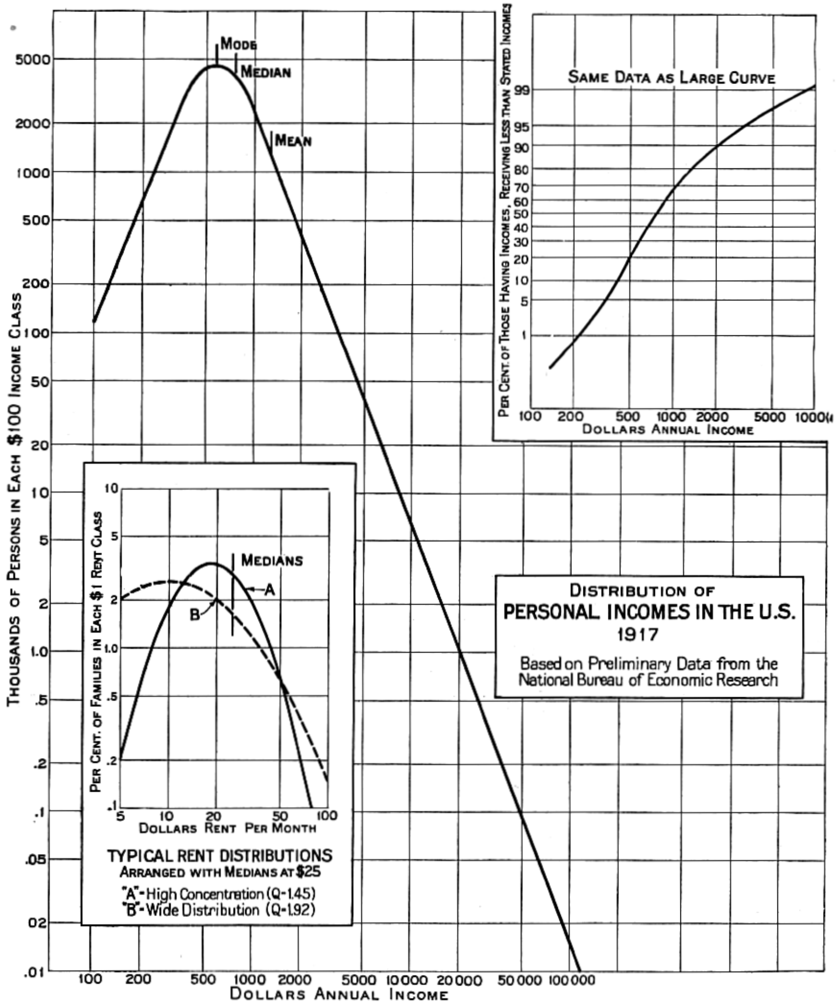

Fig. 6

Fig. 6. From the form of this curve it may be concluded that the distribution of the lower two-thirds of both incomes and rents is similar but that the spread of the higher incomes is much greater than the spread of the corresponding rents.

Another comparison of incomes and rents may be made from a second insert on the same chart. It may be readily demonstrated that the normal curve of error plots as a parabola on semi-logarithmic paper and the logarithmic skew curve as a parabola on double logarithmic paper. Two parabolas which represent extreme conditions of spread and of concentration of rents in large cities are shown. If the degree of dispersion remains fixed a change in the rent level merely shifts the parabola on the chart without changing its shape. The parabolic shape of rent curves and the hyperbolic shape of the income curve indicate that rents are somewhat less concentrated locally about their mode,[8] but are more concentrated as an entire group than are incomes. These curves can not well be superposed for comparison since areas are not equivalent on different parts of the chart. Those incomes which are closely grouped around the mode represent wage-earners of such a type that several may come from a single family. Conclusions regarding comparison of incomes and rents must be made with caution since rents are on a family basis and incomes on an individual basis. No satisfactory data are available to show the variation in income distribution between small subdivisions of the United States, as cities, but it is reasonable to assume that there is some such variation for incomes as well as for rents, although perhaps not of so great a range.

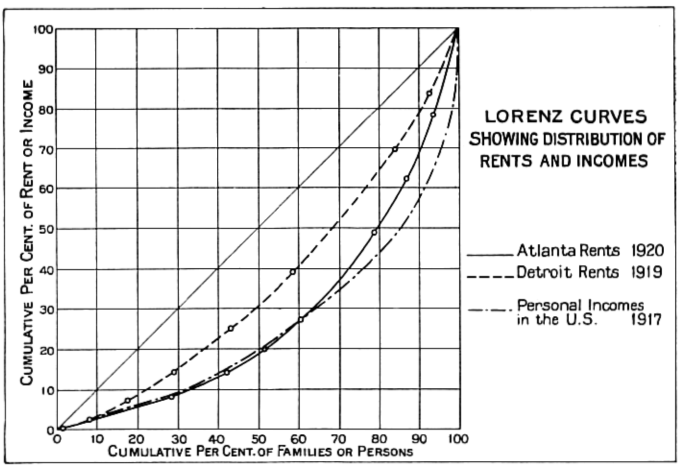

Fig. 7

A third comparison of incomes and rents is made possible by the use of Lorenz curves illustrated in Fig. 7. On this form of chart a

[8] See Appendix.

diagonal line at 45° represents a uniform distribution and the further a given curve falls to the right of and below that line the more unequal the distribution represented. If incomes were plotted on a family basis, the resulting curve would lie somewhat closer to the diagonal line than the one shown, but it is fairly evident that incomes are more unequally distributed than rents. For instance, the top 10 per cent of incomes are on the average about 42 per cent of the total income, while the top 10 per cent of rents are from 22 to 32 per cent of the aggregate rent in most cities. These three comparisons confirm the idea discussed in the first part of this paper, that the proportion of income spent for rent is less among the larger incomes.

Of the extensive data on income distributions few can well be used for comparison with rent data. In order that a cumulative curve really mean anything, it must represent an entire group, not merely items from one end or the other of the complete scale. Therefore, the various tables of earnings of working class families and individuals are of doubtful use here, although they do show, plotted on double logarithmic or logarithmic probability paper, that the type of distribution of earnings about an average value is practically identical for various nationalities in similar industries, or for men, women and children in all industries. However, the average earnings of the various classes are widely different. A few examples are shown in Fig. 5.

Income tax returns are of some interest although they are defective in several respects: they only include the upper part of society, a large number of persons fail to make returns and large amounts of income are tax exempt. Federal income tax data, which are available on a uniform basis for the years 1917–1920, may best be studied by plotting on double logarithmic paper, preferably after reducing the figures for the various states to a basis of returns per 1000 population. There are small changes from year to year in the position of the curve for any given state, which are not significant, since they may be due either to changes in the average income, or to increased efficiency of tax collection. Changes from year to year in the slope of the curve for any one state are small, indicating that there exists in each state a definite type of distribution of wealth and earning power. Differences in the position and slope of the curves for different states are conspicuous, indicating that both the per capita income and the distribution of the total income among individuals are different in different states. New York, for instance, shows a wide spread, *i.e.*, a relatively large number of very high incomes, and Iowa shows a narrow spread, *i.e.*, a large number of incomes around

\$2000–5000, and comparatively few incomes over \$20,000. New Jersey occupies an intermediate position. Alabama, more or less typical of Southern states, shows a much smaller number of returns in proportion to population than any of these states, and a distribution nearly, but not quite, as closely grouped as Iowa. The fact that a particular shape of curve is typical of a given state, and that the curves are different for different states, corresponds to similar characteristics of rent curves for cities. British income statistics show about the same degree of dispersion as do returns for the United States as a whole.

Distributions of the logarithmic skew type may be found in other fields than those of incomes and house rents. The theory has been advanced by some statisticians that while the normal curve of error is characteristic of observational errors, errors of estimate agree with that law if logarithms rather than actual estimates be considered. Price fluctuations, corporation earnings, and the profits of farmers are distributed in a similar manner. The lengths of life of telephone contracts agree quite closely with this type of distribution, if we allow for the fact that very long lives are relatively few in number because they started when the telephone business was comparatively small. A peculiarity of rent distribution is that if we choose only families having telephone service, or families having any one class of service, we obtain a logarithmic skew distribution about as closely as though we plotted all families in a city.

Application to Survey Data. The charting method described above was applied to rent data for 57 cities, both for composites of private residences, flats and apartments and for private residences alone. Table VIII gives the median rents and values of the rent dispersion index Q[9] for the composite data. Results for private residences differ in most cases only very slightly from the results given; there is no dominant tendency for the spread of private residence rents to be greater or less than that of all rents in a city, but the median rent in private residences is usually somewhat greater than that for the composite.

An effort was made to determine the significance of the various values of the index Q, but the results are chiefly negative. There is some tendency for the smaller cities to have a wide spread of rent values; *i.e.*, a high value for Q, but there is considerable scattering of the data. This tendency is most apparent in the South, where the smaller cities have extremely high values for Q. The relationship between the index Q and the per cent of families with telephone

[9] See Appendix for a quantitative definition of Q.

service is not very well defined. Cities with a very high residence development have low values for the index and Southern cities with poor residence development have high values for Q, but the intermediate scattering of data is quite wide. It might be supposed that cities with high values for Q, which indicate a wide spread of social strata, would have a relatively large number of business firms, either total or retail, to meet the widely divergent needs of the population. As a matter of fact, no such relationship is apparent. There is, however, positive correlation between Q and the proportion of institutions to population. This may be due in part to the fact that high values of Q are found in Southern cities which have separate churches and schools for whites and negroes.

Although no special significance has been found for the particular degree of rent dispersion found in any city, some interest attaches to the fact that this index remains practically constant in a given city, regardless of changes in the level of prices. The diagrams illustrating this point have already been discussed. If the type of distribution is not found constant in a particular city, it would seem probable that a change in character of the population is taking place, but a change in the average economic grade might occur without any change in the type of distribution. When two distributions, each of which agrees with a logarithmic skew curve, are added together, the new combined distribution may be represented by another logarithmic skew curve only in case both the medians and coefficients of dispersion for the two original curves are identical. It follows that if the index of rent dispersion in a city is found to be the same in successive surveys and if it may be assumed to have remained constant during the interval, then the new families which have come into a city at any time comprise a group having substantially the same coefficient of dispersion and median rent as the families which made up the original population. The apparent permanence of the type of rent distribution in a city may be considered, along with the telephone habit, as a reasonable explanation of the rather high degree of stability of station distribution by classes of service among residence subscribers.

In commercial survey work a city is divided into *market areas*, known also as *homogeneous sections*, which are so laid out that in any one section the families at any stated rent are similar telephone prospects. A study of rent distributions in market areas was carried out in a number of cities, considering only those market areas in each city which had fairly large populations. There appears to be no relationship between the index of rent dispersion and either the median rent, the per cent of families in private residences, or the per cent of families with telephone service, in the various areas in any one city. Whether a particular area is suburban or downtown, likewise has no apparent effect on the value of Q. It was found in Atlanta, where the division of the city into market areas was substantially the same in successive surveys, that the distribution index which had previously been found to be stable for cities as a whole, behaved in the same way in separate sections of the city. It was found that the rent distribution index for any single market area is smaller, usually much smaller, than the index for the entire city in which the area is located. One section in Atlanta is the only exception found to this rule. In market areas it was noted that a considerable number of the graphs on logarithmic probability paper were formed of two intersecting straight lines. This indicates that the sections are not really homogeneous, but contain elements of population radically different in character. This condition can not be obviated by the most careful laying out of section boundaries in case there exists a mixture of families of essentially different types, as when negro residences are scattered among a predominantly white population.

TABLE VIII

Indices of Rent Distribution in Large Cities

Composites of Private Residences, Flats and Apartments

	Year	Per Cent Families in Private Residences	Per Cent Families with Service	Median Rent	Q = Upper Quartile ÷ Median
New England and Eastern					
Washington	1922	66.9	43.0	$35.00	1.71
Pittsburgh	1922	61.1	37.4	28.50	1.54
Baltimore	1914	68.6	16.4	13.50	1.51
New Haven	1919	24.6	24.5	21.00	1.44
Portland, Me.	1921	36.8	49.5	23.80	1.40
Hartford	1915	20.1	25.8	19.00	1.37
Providence	1916	26.6	26.5	14.60	1.34
Springfield, Mass.	1921	29.2	45.8	30.60	1.34
Bridgeport	1920	24.2	21.4	26.50	1.32
Philadelphia	1917	81.6	18.4	17.00	1.32
Altoona	1922	90.2	45.8	24.00	1.27
AVERAGE		48.3			1.41
Central					
Chicago	1920	22.4	50.0	27.00	1.55
Cleveland	1921	45.3	32.4	35.50	1.55
Evansville	1916	89.0	34.6	12.00	1.54
Grand Rapids	1915	68.4	33.4	12.80	1.52
Milwaukee	1921	40.0	39.6	25.00	1.52
Indianapolis	1920	84.4	53.8	21.00	1.51
Akron	1920	73.0	19.3	34.00	1.41
Detroit	1919	49.9	30.6	32.00	1.40
Youngstown	1919	83.0	40.8	27.00	1.37
Toledo	1920	76.6	42.0	26.00	1.35
AVERAGE		63.2			1.47

TABLE VIII—*Continued*

	Year	Per Cent Families in Private Residences	Per Cent Families with Service	Median Rent	Q = Upper Quartile ÷ Median
Southern					
Montgomery	1913	93.6	23.7	5.50	2.82
Macon	1913	94.6	21.8	6.00	2.41
Charlotte	1914	95.8	24.6	8.00	2.30
Savannah	1916	63.5	16.2	7.80	1.98
Birmingham	1920	94.6	17.8	13.50	1.96
Memphis	1915	86.4	18.5	10.50	1.90
Atlanta	1920	83.3	28.6	19.00	1.89
Mobile	1918	92.4	19.7	7.50	1.87
Chattanooga	1915	89.2	23.7	9.00	1.83
Richmond	1922	51.6	36.5	20.00	1.75
Jacksonville	1919	75.2	24.9	13.00	1.75
Louisville	1920	71.0	28.6	14.50	1.65
New Orleans	1916	85.3	11.6	12.00	1.46
AVERAGE		83.3			1.97
Southwestern					
Tulsa	1919	88.5	40.8	$33.00	1.85
Fort Worth	1921	92.0	37.3	24.00	1.82
Little Rock	1920	93.5	40.1	19.00	1.74
Houston	1921	87.8	44.3	23.40	1.67
San Antonio	1921	91.3	31.5	20.50	1.62
Dallas	1921	88.1	54.4	38.00	1.60
Kansas City	1916	71.4	32.8	16.80	1.52
St. Louis	1917	35.3	25.0	15.00	1.50
St. Joseph	1916	90.0	41.0	13.80	1.49
Oklahoma City	1918	85.2	46.4	23.00	1.43
AVERAGE		82.3			1.64
Northwestern					
Omaha	1921	82.8	69.4	33.40	1.54
Sioux City	1918	87.6	51.0	21.50	1.53
Des Moines	1916	85.8	48.8	18.00	1.50
Duluth	1915	62.0	52.0	18.00	1.48
Lincoln	1919	87.8	60.0	23.80	1.48
Minneapolis	1921	51.6	57.2	31.00	1.45
AVERAGE		76.3			1.49
Pacific and Mountain States					
Butte	1914	75.0	35.2	17.00	1.50
Portland	1916	80.8	49.0	13.80	1.48
Denver	1917	83.7	42.4	16.00	1.47
Seattle	1918	75.7	46.8	22.00	1.46
San Diego	1918	81.8	42.4	16.00	1.44
Salt Lake City	1917	86.0	51.2	18.50	1.43
Los Angeles	1917	78.4	47.3	18.50	1.41
San Francisco	1918	53.1	47.4	21.50	1.40
Spokane	1921	86.0	57.0	23.00	1.39
Sacramento	1918	62.6	46.7	19.00	1.39
Tacoma	1921	87.8	46.3	24.00	1.35
AVERAGE		77.4			1.43

APPENDIX

Mathematics of the Logarithmic Skew Curve

Frequency curves may be symmetrical or skew. The particular symmetrical distribution known as the normal curve of error is typical of distributions of observational errors and in general of all phenomena obeying the laws of chance. It is approximated by a number of other distributions which have not obviously originated in the same way, which implies that "the variable is the sum of a large number of elements each of which can take the values 0 and 1, these values occurring independently and with equal frequency." Skew distributions may take a variety of forms but the type shown in the diagram is closely approached by a large number of rent distributions. The essential characteristic of this curve, which may be called the *logarithmic skew curve*, is that logarithms of the values of the variable are distributed according to the normal curve of error. This skew curve is of course not the only one which might be selected to represent rent data, but it presents the fewest mathematical difficulties and gives a sufficiently close approximation for all practical purposes.

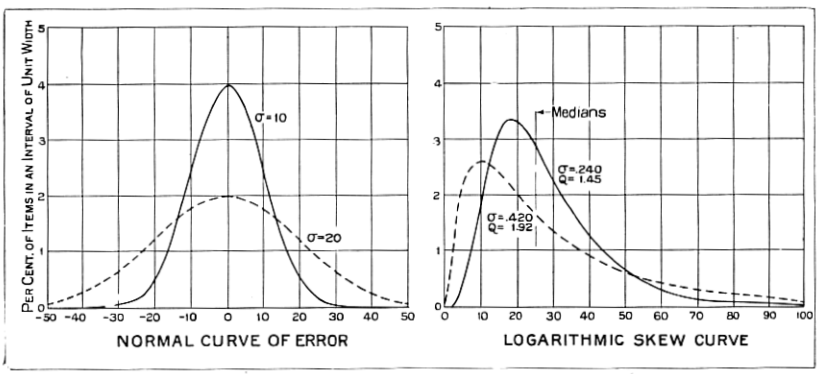

Fig. 8

The normal curve of error has the equation:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \qquad (1)$$

where e is 2.7183, the base of natural logarithms and σ is a measure of dispersion, known as the standard deviation. An ordinate of the curve is called the frequency, and expresses the fraction of the whole number of items which occurs per unit interval of the variable x.

Substituting log X for x, a new equation may be obtained, in which y is the frequency per unit of x (or log X). An expression for the frequency per unit of X is desired, which may be called Y. It may be shown that the desired equation is

$$Y = \frac{1}{X\sigma\sqrt{2\pi}} e^{-\frac{(\log X)^2}{2\sigma^2}}. \tag{2}$$

This is the equation of what we have called above the *logarithmic skew curve*, which is really not a curve of error in the same sense as equation (1) is.

In the course of this discussion it will be convenient to refer to certain features of the frequency curves by the accustomed terminology of statistics. The median item of a group is such that one-half of all the items are larger, and one-half are smaller, and is the central item when they are arrayed in order of size. The quartiles, upper and lower, together with the median, divide the array into four parts, each containing one-fourth of the items. The percentiles divide the array into 100 equal parts. The mode is that value of the variable which is of most frequent occurrence.

In the normal curve of error σ, the standard deviation, is technically defined as the square root of the mean of the squares of the deviations of the items from their mean. For present purposes it may be regarded as a measure of dispersion approximately equal to the difference between the values of x at the 84th percentile and the median. In the logarithmic skew curve σ is the difference between the corresponding logarithms.

The origin of x in the normal curve of error is the arithmetic mean, median, or mode, which are coincident. When a logarithmic scale of abscissas is introduced, the median value of x (or log X) corresponds to the median value of X, which is smaller than the mean value of X, and larger than the mode. In a logarithmic skew curve the median may be considered the origin, and at this point x (or log X) is equal to zero, and X is equal to unity. When this curve is applied to house rents the median rent occurs at this point. The relation between rents and values of X is a simple one. If rents be denoted by R, and if M be the median rent, then

$$R = MX. \tag{3}$$

The relationship of the various scales is presented in Fig. 9. The scales for X and log X may be considered fixed, and the scale for

R a movable one, as on a slide rule, corresponding values always being opposite each other.

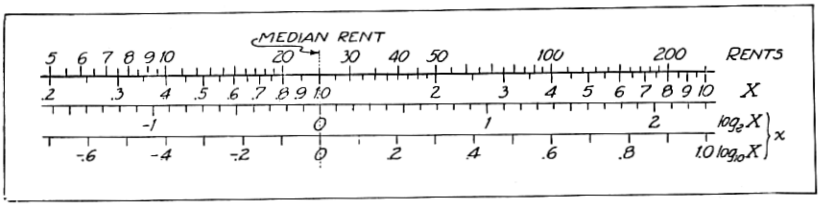

Fig. 9

Equation (2) above for the logarithmic skew curve gives the frequency per unit of X. The frequency, per unit of rent when expressed in dollars, is $1/M$ times that value, and substituting for X from equation (3), there results

$$\frac{Y}{M} = \frac{1}{R\sigma\sqrt{2\pi}} e^{-\frac{(\log R/M)^2}{2\sigma^2}}. \tag{4}$$

If it is desired to make computations from this equation, it is best to use the base 10 for logarithms rather than the natural base. For this purpose the equation becomes approximately

$$\frac{Y}{M} = \frac{.1733}{R\,\sigma_{10}} 10^{-\frac{.2171}{\sigma_{10}^2}(\log_{10} R/M)^2}. \tag{5}$$

For convenience it may be set down that

$$\sigma_{10} = 0.4343\,\sigma_e,$$

and

$$\sigma_e = 2.3026\,\sigma_{10},$$

although it will very rarely be necessary to make such computations.

It has been stated above that the median value of X (or rents expressed in dollars) may be logically regarded as the origin of the logarithmic skew curve, although X is not equal to zero at this point, but is equal to 1. If some of the other forms of statistical averages are also known, the properties of the curve may be better understood. To determine the mode, the first derivative of equation (2) of the curve is equated to zero, and there results

$$\log X = -\sigma^2,$$

or

$$X = e^{-\sigma^2}, \tag{6}$$

$$= 10^{-2.3026\sigma_{10}^2}.$$

Equation 6 above defines the peak of the curve, or the mode of the variable.

The arithmetic mean for a distribution agreeing with the logarithmic skew curve probably can not be defined by any mathematical expression sufficiently simple for practical use. It is a function of σ, but its exact position on the curve has not been determined. As applied to rent data, the mean may be computed direct from a house count summary with an error of two or three per cent. Thus found, its position on the curve is in the neighborhood of the 65th to 70th percentile.

The geometric mean coincides with the median for a logarithmic skew distribution. This follows from the fact that the median value of X corresponds to the median, which is also the mean, value of log X.

The measure of dispersion for a logarithmic skew curve is also a measure of skewness. Up to this point σ has been used as the measure of dispersion, in agreement with conventional usage. For practical purposes another measure may be substituted, which has a more readily understood meaning. This is the *quartile deviation*, known also by the misleading term *probable error*. The quartile deviation for a logarithmic skew curve is that deviation either above or below the median which includes one-fourth of all the items in the array. It may, like σ, be measured in logarithms, and

$$\text{Quartile Deviation} = 0.6745\ \sigma.$$

Perhaps the easiest mathematical conception of a measure of skewness and dispersion is that of the ratio of the upper quartile to the median. This is identical with the ratio of the median to the lower quartile, and is the number whose logarithm is the quartile deviation as defined above. We shall let this ratio be denoted by Q.

Either σ or the quartile deviation for a given set of data may be best determined from a straight line graph on logarithmic probability paper. The following table gives the positions in the array for certain convenient multiples of σ and the quartile deviation, when measured in logarithms.

Deviation from Median	*Percentile Position*
σ	84.13
2σ	97.72
3σ	99.865
Quartile Deviation	75.0
2 Qu. Dev.	91.13
3 Qu. Dev.	97.85
4 Qu. Dev.	99.65

To obtain the value of the ratio Q, take the antilogarithm of the quartile deviation determined in this manner. For ordinary purposes it is sufficiently accurate to obtain Q as the ratio of the upper quartile to the median, read from the 75th and 50th per cent lines on the graph.