Volume Title: The Behavior of Prices

Volume Author/Editor: Frederick C. Mills

Volume Publisher: NBER

Volume ISBN: 0-87014-010-8

Volume URL: http://www.nber.org/books/mill27-1

Publication Date: 1927

Chapter Title: On the Characteristics of the Population of Prices

Chapter Author: Frederick C. Mills

Chapter URL: http://www.nber.org/chapters/c5332

Chapter pages in book: (p. 311 - 355)

The month-to-month measures are so greatly affected by irregular movements in the prices of individual commodities that their usefulness appears to be limited. Indexes of displacement over twelve-month intervals are free from this defect, and provide the most satisfactory measures of changes in the internal relations among commodity prices.

VI On the Characteristics of the Population of Prices

1. The Description of Populations in Terms of Group Characteristics: Criteria of Curve Type

In describing the behavior of commodity prices in combination in the preceding sections of this chapter we have dealt with two aspects of frequency distributions of price relatives, the central tendency and the dispersion, and have in addition sought to measure shifts in the relative positions of price relatives. We come now to certain other group measures, the explanation of which calls for a brief general account of methods of describing the attributes of populations.

We were concerned, in the first chapter, with the measurement of certain characteristics of individual prices. Our interest now is in the characteristics of *populations* of prices. The difference in viewpoint is a fairly obvious one, yet one which must be stressed. The entity to be described is no longer an individual but a group, and the group has attributes of an order quite different from those of individuals. Measures of group characteristics are of two classes, those which define the type of population to which the group belongs, and those which describe attributes which are unrelated to type. In the second class are two measures which have been previously dealt with, measures of central tendency and dispersion. These describe important characteristics of frequency distributions, but characteristics in respect to which populations may differ materially while still being of the same common type. Thus a group of men might differ materially from a group of women in mean height and in the degree of variation from the mean, yet the populations of which the two groups were samples might both be of the normal (i. e. Gaussian) type in respect to height.

Differences in population type possess a degree of significance not yet fully determined, but certain matters connected with this subject are of immediate interest to the student of prices. In defining such types, when the populations are represented by frequency distributions, an obvious starting point is furnished by the normal distribution. The forces which produce this type of distribution are in some degree understood, and its basic attributes have been described. But this is by no means the only form which frequency statistics take. A number of variant types have been defined, by the use of different methods. In the present study we shall make use of the methods developed by Karl Pearson, and shall employ the criteria of curve type which he designates β_1 and β_2.[1] These criteria indicate whether a given distribution is of the normal type or of some one of the chief variant types in the Pearsonian family. In defining the distribution type, the criteria yield useful information concerning the population from which the group is drawn.

These criteria furnish the units in which Figures 42, 43, 47 and 48 are scaled, and they may be most readily interpreted in connection with these charts. The values of β_1 and β_2,* in a specific case, locate a single point, and it is the position of this point which is of importance in studying the characteristics of a given population. If a distribution is of the normal type the values of the criteria β_1 and β_2 are 0 and 3, respectively. A single point on the chart, marked G (Gaussian), represents the normal type. Any deviation from these values indicates some degree of departure from the normal distribution.[2] There is, it is obvious, an infinite number of points representing non-Gaussian distributions. In the first place, therefore, the chart affords a proper perspective in viewing the relation between the normal type of distribution and other types.

If a given distribution is not symmetrical the value of β_1 will be other than zero, the sign indicating whether the skewness is positive or negative. If a given distribution is more peaked than a normal distribution with the same standard deviation the value of β_2 will be greater than 3, while if the distribution is less peaked than the normal β_2 will be less than 3. There are thus two main lines of departure from the normal type, departures in the matter of skewness and in the matter of peakedness (kurtosis). A given distribution

[1] β_1 and β_2 are derived from the second, third and fourth moments about the mean. The formulas are given above, on p. 229.

*These two criteria are positive, by construction, but in the several charts shown in this section β_1 has been given the sign of the skewness. In preparing these charts the familiar model has been modified to facilitate the following of chronological changes in the form of price distributions. A left half, which is symmetrical with the right half as designed by Pearson and Rhind, has been added. The axis representing symmetrical distributions thus stands at the center of the chart instead of at the left, as in Rhind's diagram.

[2] A departure within certain limits may, of course, be due to errors of sampling.

may depart from the normal type in either of these respects, or in both.

A more realistic impression of the significance of changes in these criteria may be gained by an inspection of the various charts in the present volume which portray frequency distributions secured in practice. Distributions of the normal type and a variety of divergent forms may be found among those pictured in Figure 21 and in the charts in Chapter IV. A general survey of these graphs will afford concrete evidence of the differences between populations in respect to those group characteristics which are our present concern.

§ Types of Frequency Distributions

The various distribution types defined by the criteria β_1 and β_2 have been classified under several heads.[1] Symmetrical distributions, represented by points on the central axis of the chart, are of three general types. These include a symmetrical, flat-topped curve ($\beta_1 = 0$, $\beta_2 < 3$), marked as Type II[2], a symmetrical curve more peaked than the normal ($\beta_1 = 0$, $\beta_2 > 3$), marked as Type VII, and the normal or Gaussian curve ($\beta_1 = 0$, $\beta_2 = 3$), represented by the point G on the chart. There may be, of course, an infinite number of such symmetrical distributions.

The classification of skew curves is somewhat more elaborate. The three main classes, I, IV, and VI, shade off into each other, the points and lines of transition representing distinct types. The two most important of these transition forms are those of Type III and Type V. The first of these is represented by points on the line marking the boundary between curves of Type I and Type VI. Type V is a transition form between Types VI and IV. In addition to these transition types the three important sub-divisions of Type I should be recognized. These are I_I, J_I, or J-shaped curves, and U_I or U-shaped curves. (For high values of β_1 and β_2 the field of the J-curves includes curves of Type III and VI.)

Although the area representing Types I, IV and VI covers all the space within which possible values of β_1 and β_2 will fall, Pearson's system of frequency curves may be applied only within part of this field. The limit of the area within which Pearson's curves may be fitted is marked by the line $8\beta_2 - 15\beta_1 - 36$. Areas marked IV and VI (as well as the lines for curves of Types V and VII) extend below this critical line, but distributions represented by points below this line are classed as *hetero-*

[1]For a detailed description of these types Pearson's original memoirs should be consulted. The latest and most elaborate classification appears in "Mathematical Contributions to the Theory of Evolution. Second Supplement to a Memoir on Skew Variation," *Phil. Trans. Royal Society of London*, Vol. 216A (1915-16). This classification includes a number of transition frequency types not mentioned in the present brief discussion of these curves.

[2]Values of β_2 less than 1.8 (β_1 being zero) define a symmetrical U-shaped curve, classed as a subdivision of Type II.

typic. Heterotypic distributions cannot be described by any of Pearson's curve types, as derived by integration from his fundamental differential equation in its original form.[1] The reason for the existence of this heterotypic area is found in the high probable errors of the higher moments of frequency distributions. For distributions represented by points in the heterotypic area the eighth moment is infinitely large. (This is true, it should be noted, of the ideal distribution of which the actual distribution is assumed to be a sample. The eighth moment of the actual distribution would, of course, be finite.) The eighth moment enters into the probable error of the fourth moment and the probable error of β_2, hence an infinite eighth moment would render the probable errors of these quantities infinitely large. Thus if a sample distribution is represented by a point in the heterotypic area, we have no surety that if a second sample were taken from the same population approximately the same value of β_2 would be secured. A second sample might yield a distribution widely different in all its essential characteristics.

The characteristics of the chief curve types may be noted. Type I is a skew curve of limited range in both directions; Type VI is a skew curve of limited range in one direction and of unlimited range in the other; Type IV is a skew curve of unlimited range in both directions. Types III and V have the characteristics of Type VI in respect to skewness and range. Of the symmetrical types, the normal curve and curves of Type VII are characterized by unlimited range in both directions, while Type II is of limited range in both directions.

The different subdivisions of Type I are of interest in the present study, because a number of price distributions exemplify each of these sub-types. Curves of Type I_I are unimodal curves of the common pattern. J-curves have no true mode, distinct from the asymptote. For such distributions the measure of skewness (which is the ratio to the standard deviation of the distance from mean to mode) loses its customary meaning, and is not comparable to measures of skewness relating to modal curves. For curves of the U-type the value corresponding to the mode is the anti-mode, and the skewness is the ratio to the standard deviation of the distance between mean and anti-mode.

Reference to Figure 42 may make this argument clearer. There have been plotted on this chart points representing a variety of distribution types.[2] This diagram will illustrate the subject at

[1]This equation is

$$\frac{1}{y}\frac{dy}{dx} = \frac{x-a}{c_0+c_1x+c_2x^2}$$

Pearson has suggested modifications of this basic equation from which might be derived equations appropriate to heterotypic distributions. He states that the theory of curves of this type has not been published because of his failure to find any definitely homogeneous data by which it could be effectively illustrated. (See *Phil. Trans. Royal Society of London*, Vol. 216A, p. 430.)

[2]The distributions are listed below in three main classes. In recording the criteria of curve type β_1 has been given the sign of the skewness.

(Footnote continued on next page.)

present under discussion and will serve, also, as a basis for comparison when the distributions of measures relating to commodity prices are discussed.

POPULATION TYPES

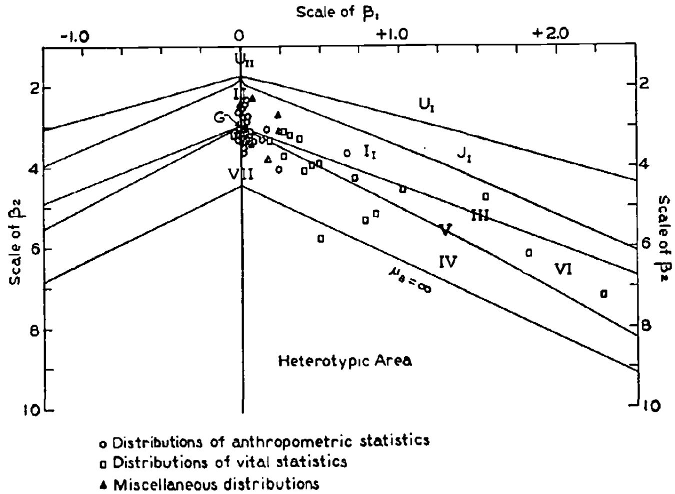
Authority	Reference	Distribution of	β_1	β_2
	Distributions of anthropometrical data			
Baxter-Pearson	*Phil. Trans. Royal Soc.*, Vol. 186-A	Heights of recruits U. S. Army	—.0058	3.0248
Davenport, C. B.	Pearson's *Tables for Statisticians and Biometricians*	Percentage of black in skin color of white and negro crosses	.6783	3.7342
Fawcett, C. D.,	*Biometrika*, Vol. I	Naqada skulls		
		Length, male	—.0046	2.6560
		female	.0139	3.1054
		Breadth, male	.0824	3.3642
		female	.0002	2.5984
		Cephalic index		
		male	.0185	2.5114
		female	.0260	2.9322
MacDonell, W. R.	*Biometrika*, Vol. 1	Head breadth of criminals	—.0104	3.0326
		Height of criminals	—.0026	3.1751
Pearl, R.	*Biometrika*, Vol. IV.	Brain weights		
		Swede, male	.0287	2.7964
		female	.0508	3.1031
		Hessian, male	.1312	3.3682
		female	.0092	2.8372
		Bohemian, male	.1623	3.1396
		female	.0130	2.8671
		Bavarian, male,	.0528	3.3026
		female	.0124	3.5408
		Bavarian, young female	.2462	4.0820
Porter-Pearson	*Phil. Trans. Royal Soc.* Vol. 186-A	Heights of St. Louis school children aged 8	—.0124	3.2350
Ranke-Pearson	*Phil. Trans. Royal Soc.* Vol. 186-A	Cephalic index of Bavarian skulls	.0079	3.6496
	Distributions of vital statistics			
Elderton, W. P.	*Frequency Curves and Correlation*	2162 Females classified according to age at death	.4950	3.9961
		Marriage rates of spinsters, classified by age	1.5572	4.7813
		Endowment insurance reserves, grouped according to ages of insured	.0000	2.9210
Pearl, R.	*Medical Biometry and Statistics*	Infant mortality rates		
		Cities over		
		25,000 1915	.2723	3.1454
		1916	—.0307	3.2440
		1917	.4033	4.1274
		1918	.3168	3.2706
		Cities under		
		25,000 1915	.1804	3.4194
		1916	.2760	3.7246
		1917	1.0243	4.6218
		1918	1.8302	6.2116
		Rural counties, total		
		1915	.8579	5.1995
		1916	.3789	3.3623
		1917	.7939	5.3916
		1918	.7391	4.2810
		Rural counties, white 1917	.4937	5.8472
		1918	.4444	3.9986
		Rural counties, colored 1918	2.2973	7.2386

(Footnote concluded on next page.)

FIGURE 42

POPULATION TYPES.

Diagram Showing the Location of Points Representing Various Types of Distributions.



There is a clear tendency for these points to cluster about the axis of symmetry, particularly about the point representing the

POPULATION TYPES (Cont.)

Authority	Reference	Distribution of	β_1	β_2
		Miscellaneous distributions		
Fisher, R. A.	*Statistical Methods for Research Workers*	Yearly rainfall at Rothamsted	.2430	2.710
Maynard	*Biometrika*, Vol. III	Number of bands on capsules of Shirley poppies	.0196	3.308
Mills, F. C.	*Statistical Methods*	Telephone subscribers classified by calls per year	.0153	3.2007
		London-N. Y. exchange rates	.0732	2.3692
Pearl, R.	*Biometrika*, Vol. V	Length of *Chilomonas paramecium*		
		Series A	.0000	2.5657
		Series B	.0311	3.1950
		Breadth of *Chilomonas paramecium*		
		Series A	.0011	3.1332
		Series B	.1852	3.8831
Surface, Frank M.	*Biometrika*, Vol. III	Size of litters Poland China sows	.0284	3.3622
		Size of litters Duroc Jersey sows	.0167	3.3422
Venn-Pearson	*Phil. Trans. Royal Soc.* Vol. 186-A	Barometric heights	.2440	3.1739
Weldon-Pearson	*Phil. Trans. Royal Soc.* Vol. 186-A	Right antero-lateral margin of female Naples crabs	.0267	3.1128

normal distribution. Yet there are pronounced departures from this Gaussian point, departures which, in the present case, are mostly in the direction of positive skewness. The points representing anthropometric distributions are, as a group, concentrated most closely about the Gaussian point. Distributions of infant mortality rates shows the widest departures from that point.

It is to be noted that, while there is some variation between the results secured from different samples relating to a given characteristic (such as brain weights), the range of variation is small for most of the groups here represented. There is one rather conspicuous exception to this rule. The criteria of curve type vary materially for different samples of infant mortality rates. Reference to the footnote describing the distributions will show that there were marked differences between the populations from which the various infant mortality samples are drawn, a fact which undoubtedly accounts for a large part of the variation observed.

The criteria β_1 and β_2 sum up all the main characteristics of a frequency distribution, other than those defined by the measures of central tendency and dispersion. Specifically, information as to the following features of a distribution is given by these criteria:

a. The direction and degree of skewness. Any departure from symmetry is shown by a movement of the Beta point[1] away from the central axis at which the value of β_1 is zero. If the skewness is positive the point will lie to the right of the axis of symmetry, if negative, to the left of that axis, as it appears in these charts.
b. The shape of the distribution in respect to kurtosis. If the concentration at the central value is greater than would be found in a normal distribution having the same standard deviation as the distribution in question (i. e. if the distribution is more peaked than the normal) the point will lie below a horizontal line passing through the axis of symmetry at the point $\beta_2=3$. If the distribution is less peaked than the normal the point relating to it will lie above this line.
c. The "ideal" frequency type to which the distribution corresponds. The most important of these types have been discussed above.

[1]The term *Beta point* has been used for convenience in referring to any point the coordinates of which are defined by values of β_1 and β_2.

There is a further type of information which these criteria may yield, though the present state of our knowledge on the subject to be discussed will hardly justify broad generalizations or unqualified conclusions. This information relates to the inherent stability of given distributions, and of the populations from which they are drawn.

Reference was made in an earlier section to the problem of price stability. External stability, or stability of the price level, was distinguished from internal stability. The latter term was used to describe stability of established relationships between individual prices and groups of prices. We are now dealing with a third (not unrelated) type of stability, the inherent stability of frequency distributions of price relatives.

What meaning can we attach to the term "stable frequency distribution"? Any frequency distribution represents the resultant of a number of separate forces. In the absence of correlation between the agencies at work, and assuming that there are so many forces in operation that no one force exerts a preponderant influence upon the data classified, something approaching stable equilibrium will be secured in the frequency distribution. In the perfect example, when positive and negative deviations are equally likely to occur, and when the distribution follows that curiously universal pattern first traced by Laplace and Gauss, a condition of true stability has been attained. The forces tending to produce values above the mean are precisely balanced by forces pulling in the other direction. So fine is the balancing of diverse forces that, knowing the normal law of error to be in operation, it is possible to foretell, within limits capable of precise evalution, the exact proportion of observations which will depart by given amounts from the mean value. Under these conditions, says Pearson, ". . there is production and destruction impartially around the mean."

A stable frequency distribution may be thought of, then, as one which embodies the balanced condition reflected in most perfect form in the normal curve. There is variation, but the forces of disruption are held in check by cohesive agencies. The population from which the distribution in question is drawn is not subject to "sport" types departing from the norm by amounts which are excessive, in terms of the standard deviation of that population. Sporadic variation plays, in perfect balance, within assignable limits. *Stability of type* is generated by the balance of forces responsible for individual variation. In this sense, then, stability is a group characteristic. Frequency distributions constructed from measurements of those biological characteristics which are most stable (e. g., stature, skull dimensions, brain weight) are of this normal type, or approximate it closely. Unstable organisms may be expected to yield distributions departing materially from the normal type, these departures representing break-downs, "sports," and erratic individual variations.

If a distribution of the normal type represents a condition of stable equilibrium, it may be assumed that departures from that type indicate the presence of some unstable element or elements. Yet all departures are not of equal significance. The Gaussian distribution is secured, Pearson has pointed out, when each contributory cause group is independent and when (if the number of groups be not very large) each cause group is of equal valency and contributes with equal frequency results in excess and defect of its mean contribution. The conditions which yield a Type III distribution are closest to those which give rise to the Gaussian type. The contributory cause groups are of equal valency and are independent, as for the Gaussian type, but Type III differs from the Gaussian in that each cause group does not necessarily contribute with equal frequency results in excess and defect of its mean contribution. Type III corresponds to the skew binomial.[1] In employing the fundamental differential equation used in deriving the other curve types in Pearson's system it is not necessary to assume the independence or equal valency of the cause groups, nor the equality of contributions above and below the mean.

In studying the characteristics of frequency distributions of price relatives it will be of interest to observe the degree of departure of given distributions from the normal type and from Type III, considering these to be representative of more stable populations than the other frequency types in the Pearsonian system.[2]

[1]See Pearson, *Tables for Statisticians and Biometricians*, p. lxi, and Pearson, "Skew Variation in Homogeneous Material," *Phil. Trans. of the Royal Society*, Vol. 186 A, pp. 356-360,381. See also, Kelley, Truman L., *Statistical Method*, p. 148, for a comment on the point in question.

[2]This subject of the relation between curve type and stability of populations has been discussed in a very illuminating fashion by Truman L. Kelley (*Statistical Method*, pp. 138-150). In the course of this discussion Kelley emphasizes certain distinctive characteristics of the distribution types which have been mentioned above as being most stable. Type III distributions, he shows, are the only ones which do not possess certain infinite positive moments (i. e. moments which become infinite for a fitted curve). Just as these Type III distributions are unique, in that all their positive moments are finite, so distributions of Type V are unique, in possessing finite negative moments. The lines representing Type III and Type V distributions meet, it may be observed, at the Gaussian point (see Figure 42). From which facts Kelley draws the conclusion:

Statements in the preceding paragraphs suggest some of the causes of instability of type, as that phrase is used in the present discussion. Unstable elements may be present when essentially homogeneous data are subject to the interaction of factors too limited in number or of such marked inequality that certain forces exert a preponderant influence. Again, correlation between the factors affecting the quantities measured tends to destroy the balance characteristic of stability. Probably the most important of the conditions which give rise to departures from Type III, as well as from the normal type, is this factor of dependence of the contributory causes or, in Kelley's phrase, of "uncompensated correlation between cause groups."

In addition to that type of instability which is evidenced by "skew variation in homogeneous material" we may have the instability which is found when the data classified consist of a mixture of heterogeneous materials. In this case the groups which are mixed may be subject, in their variation, to essentially different forces, a fact which would ordinarily give rise to distribution types departing materially from the normal.[1] In the study and comparison of price groups at a later point in the present investigation (to be discussed in a later volume) we shall return to the subject of heterogeneity of populations, and shall discuss methods of testing for heterogeneity.

As distributions depart from the normal type there is always some loss in both the significance and the reliability of the statistical measures descriptive of the populations represented. They lose in significance because there is not the clear central tendency and the uniform deviation from the central tendency which may be most accurately described. They lose in reliability because of an increase in the sampling errors to which all statistical measures are subject. As the degree of departure from the stable Gaussian type and Type III increases, the tendency toward a breakdown of the distribution type increases. The statistical evidence of this tendency

"If......finite positive moments are of most importance III is the most stable of all the types; however, should finite negative moments be of greater importance than positive, Type V would be the most stable; and if the possession of both finite positive and negative moments is material then the normal distribution is the most stable curve within all the types." (p. 143).

[1]This is not necessarily true, for a heterogeneous distribution might take the Gaussian form. In an early article in *Biometrika* ("Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson," A Rejoinder. *Biometrika*, Vol. IV) Pearson effectively demonstrated that correspondence of a distribution to the Gaussian type is not necessarily evidence of homogeneity, and that departure from the normal type is not evidence of heterogeneity.

is found in the magnitude of the sampling errors, which may become so large that there can be no confidence that a second sample from a given population will yield a distribution of the same type as the first sample.

§Criteria of Instability

Somewhat more specific information concerning instability of type is afforded by two of the criteria Pearson has employed in defining his basic frequency types. These are the measures represented by the symbols κ_1 and r. The criterion κ_1 is derived from the following relationship:[1]

$$\kappa_1 = 2\beta_2 - 3\beta_1 - 6$$

The type of information given by κ_1 may be made clear by reference to Figure 42. On this chart a line appears, representing the locus of the points for which $\kappa_1 = 0$. This is identical with the line representing Type III distributions. If κ_1 is positive, in a given case, the Beta point falls below this line, in the regions of Type IV, V and VI. If κ_1 is negative the Beta point falls above this line, in the Type I area. The value of κ_1 indicates the degree of departure of the given point from the line and, correspondingly, the degree of departure of the given distribution from Type III.

Some of the characteristics of Type III distributions have been noted above, with particular reference to the problem of stability of type. On the assumption that finite positive moments are of greatest importance in the material with which we are at present dealing, the criterion κ_1 may serve as a measure of stability. A value of zero means perfect stability, in that all positive moments are finite. As the value of κ_1 departs from zero the moments, beginning with the highest, become infinite. The greater the value of κ_1, on this assumption, the less stable is the distribution.

In deriving the other criterion which concerns us here we set

$$r = \frac{6\,(\beta_2 - \beta_1 - 1)}{2\,\beta_2 - 3\beta_1 - 6}$$

Since the denominator of the right hand member of this equation is equal to κ_1 it is clear that r and κ_1 are related. When κ_1 is equal to zero, as it is for the normal curve and for distributions of Type III, r will be equal to infinity. It descends from infinity, with positive values, below the Type III line, and descends, with negative values, above that line. The numerator of the above fraction is necessarily positive, hence r will agree in sign with κ_1. But r yields certain additional information which κ_1 does not. On a preceding page reference has been made to the high probable errors of the higher moments of frequency distributions. It has been pointed out that for distributions represented by points in the heterotypic area the eighth and all higher moments (of the ideal distribu-

[1]In this and similar equations which follow β_1 is taken as positive. This measure is positive by construction, but in preparing the graphs β_1 has been arbitrarily given the sign of the skewness.

tions represented by the sample) are infinitely large. The criterion r in dicates the order of the moments which, in a given case, would be infinite If r is equal to any value, m, then the moment of the order $m + 1$ woul be infinite. Thus if r is equal to 7, the eighth and all higher moments ar infinitely large for the ideal distribution of which the actual distributio is a sample. Under the conditions last named the value of β_2 as deter mined from the sample has an infinite probable error. We may, therefore determine from the value of r in a given case whether a distribution i heterotypic. This will be true when r, being positive, is equal to or les than 7. If r is equal to 3, the fourth and all higher moments are infinitel large and the value of β_4 for the ideal curve becomes infinite. If r has value between 0 and 1 the second moment and the standard deviatio become infinite.

Since r indicates the order of the moments which break down fo specific distributions (i. e. become infinite for the corresponding ideal dis tributions) it may appropriately be employed as a measure of stabilit of type, as that term is here used. The greater the value of r the highe is the order of the moments which become infinite and the more stabl are the distributions represented. As r approaches infinity the cor responding distributions approach the normal type, or Type III. Thes two types appear, as has been pointed out, when the causal factors ar independent and when the cause groups are of equal valency. Th lower the absolute value of r, whether the sign be positive or negative, th less stable is the distribution.[1]

One of the most important aspects of this problem of instabilit of distributions relates to the heterotypic area which is shown i Figure 42, and to which reference has been made above. No pre cise meaning has been given to the term *heterotypic.* It is clea that if a distribution falls in this area the probable errors of th criteria of curve type are so great that there can be no confidenc that other samples from the same population would yield result within reasonable limits of those secured from the first sample But of what sort of distribution is this true? It has been suggeste that a tendency toward bi-modality may account for the location o a distribution in this area, but the results of the present study sup port Truman Kelley's assertion that this is not necessarily true. I is Kelley's contention that the location of a distribution in th heterotypic area indicates the presence of essentially unstabl elements, elements conducing to a breakdown of type. The hig probable errors of the moments indicate not only the unreliabilit

[1]The uses of r are discussed in the various references to Pearson which have bee given above. Extensive use of this criterion has been made by R. A. Fisher in his m moir "On the Mathematical Foundations of Theoretical Statistics" (*Phil. Trans. of t Royal Society*, Vol. 222-A, 1921-22, pp.309-368). In defining the frequency types the Pearsonian system Fisher employes a diagramatic method which is based primari upon r, and which differs materially from the method represented by Figure 42.

of sampling, but a tendency toward erratic and disruptive movements on the part of individual members of the population. With reference to animal organisms, we might say that it indicates the possibility of thoroughly abnormal "sports" which would break completely away from the prevailing type. It suggests that the pull toward the center of gravity of the distribution may be so relaxed that individual members of the population might break loose, as planets might escape from a parent sun. This is, however, a field in which the present state of our knowledge does not justify any final conclusions. It is certain that for distributions which fall in this region the ordinary measures used in describing frequency distributions and in generalizing from them lose materially in reliability and utility. And there are some reasons for thinking that this failure of the orthodox type is due to elements of instability in the populations from which given samples are drawn.[1]

2. Characteristics of Frequency Distributions of Price Relatives

With the above considerations before us we may proceed to a discussion of the attributes of the population of prices. For the present we are concerned with just one characteristic of commodity prices—degree of change between specific dates, as measured by link and fixed base relatives—and conclusions concerning the population of prices are to be interpreted with this in mind.

In Tables XIX to XXVII, in the Appendix, are given the criteria of curve type and related measures descriptive of various distributions of price relatives for the years from 1891 to 1926. In order that chronological changes in the character of certain of these distributions may be followed, the criteria β_1 and β_2 for two sets of distributions (unweighted fixed base relatives in natural form and

[1]R. A. Fisher has explained the failure of the orthodox types within the heterotypic area as due to the inadequacy of the method of moments in estimating the parameters of populations from samples which are heterotypic (see Fisher, "On the Mathematical Foundations of Theoretical Statistics," p. 346). Pearson's system rests upon the method of moments. Moreover, Fisher contends, the inadequacy of the method of moments is not confined to dealing with heterotypic distributions; this method falls far short of a perfect standard of efficiency in estimating the form of distributions of other types, notably J-shaped distributions. The efficiency of the method of moments is high, says Fisher, only in the neighborhood of the Gaussian point.

It is probably not necessary to set in opposition these two explanations of the failure of the Pearsonian frequency types as means of estimating the form of ideal curves corresponding to certain distributions. It may be granted that the method of moments loses in efficiency as it is applied to distributions departing materially from the normal type. Yet the departure of a given distribution from that type may still be taken to mean that the population from which this sample was drawn is subject to the play of unbalanced forces and contains unstable elements.

weighted logarithms of link relatives) have been plotted in Figures 43 and 48.

The first point to be noted in an examination of these diagrams is the fact of extreme year-to-year changes in distribution type. The population with which we are dealing is one which changes its fundamental characteristics from year to year. No such extreme differences are to be found between the results secured from different samples of brain weights, infant mortality rates, or skull dimensions, measures for which were presented above. It may be, however, that we are not justified in thinking of these results for different years as samples from a single population. Link relatives measuring the changes in individual commodity prices between 1915 and 1916 reflect the play of a set of forces quite different from those which influenced the link relatives for 1921, on the 1920 base. If we are to think of price relatives as constituting a population, we should perhaps view the relatives for each year as a population by themselves. Whether the specific changes from one year to the next possess economic significance is discussed below.

The descriptive measures given in the Appendix tables permit detailed study of the populations secured by combining price relatives in different forms. Various distributions may be compared in regard to stability of type, degree of skewness and dispersion, closeness of correspondence to the normal or any other ideal frequency type, and in other respects. In such a study a three-fold comparison is possible.

Distributions of the logarithms of price relatives may be compared with distributions of price relatives in natural form.

Distributions of weighted price relatives may be compared with distributions of unweighted relatives.

Distributions of fixed base relatives may be compared with distributions of link relatives.

In each case, of course, the comparison involves distributions differing in one characteristic only. Thus in comparing logarithmic and natural distributions the measures relating to unweighted logarithms of link relatives for 1920 would be compared with similar measures relating to unweighted link relatives in natural form for 1920. Measures of dispersion, and all other measures, were put on comparable bases before the comparison was made.[1]

[1]In the case of measures of dispersion this was done by expressing $.6745\sigma$, in natural numbers, as a percentage of the mean. This value was compared with the index of dispersion derived from logarithms. (The derivation of this index is explained on p. 257.)

§Comparisons of Distributions of Price Relatives
Arithmetic and Logarithmic Distributions

The first of these comparisons is permitted by the figures in the table below.

TABLE 111

COMPARISON OF DISTRIBUTIONS OF LOGARITHMS OF PRICE RELATIVES AND DISTRIBUTIONS OF PRICE RELATIVES IN NATURAL FORM
(Based upon prices during the period 1913-1926)

Characteristic	No. of pairs of distributions compared	Result
1. Mean value		Logarithmic (geometric) mean necessarily smaller in all cases.
2. Dispersion	136	Log dispersion smaller 70 times.
3. Skewness	50	Log skewness smaller 37 times.
4. Sign of skewness	50	Log skewness negative 26 times. Natural skewness negative 6 times.
5. Kurtosis	50	Log closer to normal 36 times.
6. Location in respect to heterotypic division	50	Number of log distributions heterotypic: 30 Number of natural distributions heterotypic: 21
7. Stability (measured by κ_1)	50	Log more stable 34 times.
8. Curve type[1]	50	Logarithmic distributions: Type I 3 Type IV 43 Type VI 4 Natural distributions: Type I 8 Type IV 28 Type VI 14

[1]A number of the distributions classified in this table might have been placed among the transition types, if account were taken of the sampling errors to which the criteria β_1 and β_2 are subject. In view of the magnitude of these sampling errors for points near the heterotypic limit and within the heterotypic area, it has been considered more significant to group the distributions into the three main classes, and their sub-classes, which are defined by areas, not by points or lines.

Of the 8 natural distributions of Type I, 5 were of the sub-type I_J; of the 14 of Type VI, 8 were of the sub-type VI_J.

Of the 3 logarithmic distributions of Type I, 1 was of the sub-type I_J; of the 4 of Type VI, 1 was of the sub-type VI_J.

(See p. 314 for a brief description of these curve types.)

Only 50 pairs of distributions are directly comparable in the matter of curve type. We have data, however, relating to 96 logarithmic distributions (of which approximately two-thirds were based upon link relatives) and 94 distributions of relatives in natural form (of which approximately two-thirds were constructed from fixed base relatives). The comparison is not as trustworthy as that given in the table, but it possesses significance, since no essential difference was found between fixed base and link relatives in this regard. The classifications follow:

Curve type	Number of logarithmic distributions of given type	Number of natural distributions of given type
I	8	11
IV	85	62
VI	3	21
	96	94

One of the logarithmic distributions of Type I and one of the logarithmic distributions of Type VI belonged to sub-groups of the I_J and VI_J types, respectively.

Five of the natural distributions of Type I were of sub-type I_J, and 11 of the natural distributionsof Type VI were of sub-type VI_J.

The logarithmic and arithmetic measures of dispersion were commented upon in an earlier section. It was there pointed out that there is no evidence of a consistent tendency for one to exceed or fall below the other. When prices are seriously disturbed, however, as they were during the war and immediate post-war years, the arithmetic standard deviation is subject to extreme fluctuations. During the years 1915 to 1918 this was particularly true of unweighted distributions of price relatives on a fixed base. Though the logarithmic measures of dispersion were in many cases greater than the corresponding natural measures, they never exceeded the arithmetic measures by any considerable margin. During years when price disturbances are less extreme there appears no tendency for the logarithmic dispersion index to exceed or fall below the corresponding arithmetic measure.

The logarithmic distributions tend, on the whole, to be more symmetrical than the arithmetic. Out of the 50 comparisons which were possible, the measures of skewness of the logarithmic distributions were smaller in 37 cases than the measures relating to corresponding arithmetic distributions. This is due, of course, to the fact that the range of positive deviations is much less in the logarithmic than in the natural distributions. Price relatives in natural form are subject to extreme positive deviations, since the upper limit of such a relative is infinity, while the lower limit is zero. For the same reason we find that the logarithmic distributions were negatively skewed 26 times out of 50 cases, while natural distributions were negatively skewed in only 6 out of 50 cases.

The logarithmic distributions tend, also, to be less peaked than the arithmetic. The general tendency of all distributions of price relatives is to be more peaked than the normal, but this tendency is more pronounced in the arithmetic type.

On *a priori* grounds (reasoning from the known characteristics of price relatives in logarithmic and natural form in the matter of range) one might expect Type IV, the skew, peaked curve with theoretically unlimited range in both directions, to predominate among logarithmic distributions, and Type VI, which is skewed and peaked and which has a theoretically unlimited range in only one direction, to predominate among arithmetic distributions. In fact, both logarithmic and arithmetic distributions are preponderantly of Type IV, but the proportion of distributions of this type is much greater among the logarithmic than among the arithmetic distributions. Among the arithmetic distributions Type VI ranks next in importance to Type IV. Type I has 3 representatives among the logarithmic distributions and 8 among the natural. (The count is based upon the distributions for the years 1914-1926.)

Several measures which serve as indexes of the inherent stability of distribution types have been discussed above. Judged with reference to distance from the Type III line, which represents, presumably, the most stable skew type, logarithmic distributions of price relatives appear to be more stable than natural distributions. In 34 out of 50 cases distributions based on logarithms deviate from this type by smaller amounts than the corresponding natural distributions. (The criterion here employed is κ_1.)

Another comparison may be made, employing the criterion r, which indicates the "break-down" point of different distributions. That is, it indicates the order of moments which become infinite. The results of the comparison, which is restricted to positive values of r, are summarized in the following table. The distributions here compared relate to the period 1914-1926.

Range of positive values of r	Number of distributions for which the values of r fall within the stated limits[1]	
	Logarithmic	Natural
7.00, or below	30 out of 50 cases	21 out of 50 cases
6.00, or below	17 out of 50 cases	15 out of 50 cases
5.00, or below	13 out of 50 cases	10 out of 50 cases
4.00, or below	1 out of 50 cases	4 out of 50 cases

[1]The various sub-totals in this table are cumulative, hence the sum of the items in each column is not necessarily equal to the total number of distributions (50) studied in each case. Thus, for example, the 17 logarithmic distributions for which the value of r was 6, or below, are included in the 30 distributions for which the value of r was 7, or below. The 20 logarithmic distributions and the 29 natural distributions for which the value of r exceeded 7 are not, of course, listed in the above table.

If r has a value of 7 or less, it will be recalled, the eighth moment of the ideal distribution corresponding to the sample is infinite, and the probable error of β_2 is infinite. This marks the limit of the heterotypic area. If r has a value of 6, or less, even greater instability is indicated, for this means that the seventh and all higher moments are infinite.

Judging from all the entries but the last in the above table, logarithmic distributions appear to be less stable than natural distributions. Thirty of the 50 logarithmic distributions are heterotypic, while only 21 of the corresponding 50 natural distributions fall in this region. The last entry in the table shows that 4 of the natural distributions and 1 logarithmic distribution are characterized by infinite fifth moments.

The better showing of the natural distributions in this comparison is not due to any tendency on their part to cluster about the normal point, or about the stable Type III line, as was indicated by the values of κ_1. During times of extreme price disturbance natural distributions, particularly unweighted distributions, tend to deviate above the Type III line, into the area of J-curves of Type I. This is the type of distribution followed by unweighted fixed base relatives in natural form in 1915, 1916, 1917 and 1918, by unweighted link relatives in natural form in 1915, and by weighted link relatives in natural form in 1923. Judging solely on the evidence of positive moments these are stable distributions, and the measures describing them are valid. R. A. Fisher has shown, however, that the efficiency of the method of moments in locating and scaling such J-curves is very low. All J-curves fall beyond what Fisher calls the "region of validity" of the first moment. The upper boundary of the region of validity of the second moment is set by the line r = -4, a line beyond which 5 of the distributions of price relatives in natural form fall. The means and standard deviations relating to these distributions are not characterized by the infinite probable errors which are found when positive moments of a low order are infinite, but are equally unacceptable because of their inadequacy as measures descriptive of the population of prices.[1] Because of the existence of this upper boundary to the efficiency of the statistics descriptive of the natural distributions, it does not seem justifiable to conclude that the natural distributions are preferable to the logarithmic because a smaller proportion of the natural distributions fall in the heterotypic area. In fact, were one to base a choice between the two upon the basis of the efficiency of the first and second moments (i. e. of means and standard deviations) the logarithmic distributions would be selected. For the proportion of logarithmic distributions falling within the region of validity of the first and second moments is slightly greater than the proportion of natural distributions.

In a final comparison we may determine the number of logarithmic and natural distributions which correspond in type to the normal curve of error, noting at the same time the character of the deviation from the normal type. The results appear in the following table. In preparing this table the actual measures of skewness and kurtosis have been compared with the standard errors of these measures for samples from normal distributions. If a given measure of skewness or kurtosis differed from the Gaussian value of 0 by less than 2.576 times its standard error, the deviation was not considered significant. A greater deviation was taken to mark a significant departure from the normal type.[2] The distributions here classified relate to the period 1914-1926.

Form of price relatives	Number of distributions deviating from normal type in respect to			Number of distributions conforming to normal type in respect to skewness and kurtosis (within sampling limits indicated in the text)	Total
	skewness alone	kurtosis alone	skewness and kurtosis		
Natural	2	2	44	2	50
Logarithmic	1	8	38	3	50
Total	3	10	82	5	100

[1]This upper boundary, says Fisher, is set not because the probable errors of the moments become infinite, but because "the ratio of the probable errors of the moments to the probable error of the corresponding optimum statistics is great and tends to infinity as the size of the sample is increased." ("On the Mathematical Foundations of Theoretical Statistics," p. 347.) The optimum statistics are those of which the likelihood is greatest. The concept of maximum likelihood is developed by Fisher in the article named, and is exemplified in Fisher's *Statistical Methods for Research Workers*, (Edinburgh, 1925).

[2]The standard error multiplied by 2.576 marks the limit beyond which a deviation would fall as the result of chance less than 1 time out of 100.

Most of the distributions in both groups differ from the normal type in respect to both skewness and kurtosis. The tendency toward greater symmetry in the case of logarithmic distributions accounts for the larger number in this group which deviate from the normal type in respect to kurtosis alone. Only five of the total group of 100 distributions could be considered normal, even within the generous sampling limits upon which this classification rests.

Weighted and Unweighted Distributions

The effect of weighting upon the form of frequency distributions of price relatives is revealed by the entries in the following table.

TABLE 112

COMPARISON OF WEIGHTED AND UNWEIGHTED DISTRIBUTIONS OF PRICE RELATIVES

(Based upon prices during the period 1890-1926.)

Characteristic	No. of pairs of distributions compared	Result
1. Mean value	136	Wtd. smaller 57 times
2. Dispersion	136	Wtd. smaller 74 times
3. Skewness	95	Wtd. smaller 57 times
4. Sign of skewness	95	Wtd skewness negative 35 times Unwtd. skewness negative 33 times
5. Kurtosis	95	Wtd. closer to normal 68 times
6. Location in respect to heterotypic division	95	No. of wtd. distributions heterotypic: 41 No. of unwtd. distributions heterotypic: 49
7. Stability (measured by κ_1)	95	Wtd. more stable 61 times
8. Curve type[1]	95	Wtd. distributions Type I 10 Type IV 71 Type VI 14 Unwtd. distributions Type I 9 Type IV 76 Type VI 10

[1]Six of the 14 weighted distributions classed in Type VI were of the sub-type VIJ. Six of the 9 unweighted distributions in Type I were of the sub-type IJ, and 6 of the 10 of Type VI belonged in the sub-group VIJ.

Weighted and unweighted distributions of price relatives have been compared in respect to mean value and dispersion in earlier sections. It may be noted here that the effect of the introduction of weights based upon values marketed in the period 1919-1923 was to increase the mean value during the period 1891-1913, and to decrease the mean value during the years 1914-1926. There seems to have been no consistent difference between weighted and unweighted measures in the matter of dispersion prior to 1914. During the years of violent price change since 1914 the dispersion of weighted distributions has been smaller in about three-quarters of the cases compared.

The departures from symmetry were somewhat less pronounced for weighted than for unweighted distributions, but this result is due entirely to the situation since 1914. The weighted skewness was less in 34 out of 50 cases in this period. During the less disturbed period preceding, the skewness of the weighted distributions was less in 23 out of 45 cases. The skewness was negative in almost precisely the same proportion of unweighted and weighted distributions.

The use of weights serves substantially to decrease the peakedness of price distributions. The weighted distributions were closer to normal in this respect during both the earlier and later years of the period covered, the advantage being in favor of the weighted distributions in 68 out of 95 cases.

There is no material difference in the type of ideal distributions to which the members of the two groups belong. In both cases over 70 per cent of the distributions are of Type IV, with Type VI and Type I standing next in order of importance.

The available evidence indicates that combinations of weighted price relatives yield more stable distributions than do combinations of unweighted relatives. Weighted distributions are more stable in 61 cases out of 95, as judged by κ_1 (which measures the distance from the Type III line). When the criterion r is employed in the comparison, the following results are secured.

Range of positive values of r	Number of distributions for which the values of r fall within the stated limits	
	Unweighted	Weighted
7.00, or below	49 out of 95 cases	41 out of 95 cases
6.00, or below	40 out of 95 cases	30 out of 95 cases
5.00, or below	19 out of 95 cases	17 out of 95 cases
4.00, or below	1 out of 95 cases	8 out of 95 cases

The proportion of unweighted distributions which are heterotypic (i. e. for which r is positive and equal to 7 or less) is greater than the corresponding proportion of weighted distributions. On the other hand, a larger number of weighted distributions fall in the last group in the table, in which are tabulated the distributions for which the fifth and all higher moments are infinite.

In the following table unweighted and weighted distributions are compared in the matter of correspondence to the normal curve of error. The distributions here classified relate to the period 1891-1926.

Form of price relatives	Number of distributions deviating from normal type in respect to			Number of distributions conforming to normal type in respect to skewness and kurtosis (within the sampling limits indicated in the text)	Total
	skewness alone	kurtosis alone	skewness and kurtosis		
Unweighted	2	17	72	4	95
Weighted	4	13	64	14	95
Total	6	30	136	18	190

The most pronounced difference between the two sets of figures presented is found in the column showing the number of distributions conforming to the normal type. Fourteen weighted and only four unweighted distributions are in this group. The proportion of the total number of distributions which may be considered normal (within the sampling limits defined above) is small in both cases, but it is relatively much greater for the weighted measures.

Distributions of Fixed Base and Link Relatives

The final comparison involves distributions constructed from fixed base and link relatives. The basic data appear in the following table.

TABLE 113

COMPARISON OF DISTRIBUTIONS CONSTRUCTED FROM FIXED BASE AND LINK RELATIVES OF COMMODITY PRICES, FOR THE PERIOD 1914-1926

Characteristic	No. of pairs of distributions compared	Result
1. Mean value		(Comparison not significant)
2. Dispersion	128	Link smaller 128 times
3. Skewness	48	Link smaller 26 times
4. Sign of skewness	52	Link skewness negative 18 times Fixed base skewness negative 17 times
5. Kurtosis	48	Link closer to normal 31 times
6. Location in respect to heterotypic division	52	Number of link distributions heterotypic: 27 Number of fixed base distributions heterotypic: 27
7. Stability (measured by κ_1)	48	Link more stable 28 times
8. Curve type[1]	52	Link / Fixed base Type I: 6 / 5 Type IV: 35 / 40 Type VI: 11 / 7

[1]One of the 6 link distributions classified as of Type I was of sub-type I$_J$; 4 of the 11 classed as of Type VI belonged to sub-group VI$_J$.
The 5 fixed base distributions of Type I all were of sub-type I$_J$; 5 of the 7 classed as of Type VI belonged to sub-group VI$_J$.

It is to be expected, of course, that the dispersion of link relatives should be smaller than that of fixed base relatives. There is no material difference in the matter of skewness. It is apparent that the greater dispersion of fixed base relatives, as the base becomes further removed, does not result in any appreciable change in the degree of skewness. Nor is there an appreciable difference between the two types of distributions in respect to the direction of skewness.

The link distributions are distinctly less peaked than the fixed base distributions for the period from 1914 to 1926. In 32 out of 48 cases the link distributions are more flat-topped than the others, and in 31 of these 32 cases the link distributions are closer to the normal curve in this respect. In interpreting these results it must be remembered that the measure of peakedness refers always to a normal distribution having the same standard deviation as the distribution in question. Since the dispersion of the link distributions is less, in all cases, these distributions look more peaked when presented in graphic form. They are not so, however, when the criterion of peakedness is a normal curve with the same degree of dispersion. It is possible that a somewhat different proportion would prevail if comparison of a number of pre-war distributions were possible. The abnormally wide dispersion of fixed base relatives between 1915 and 1918 is reflected in high values of the measure of kurtosis.

In the matter of curve type there is no material difference between link and fixed base distributions. In both cases Type IV predominates.

The criterion κ_1 shows that the distributions composed of link relatives are closer to Type III in 28 out of 48 cases. If we use the values of r in testing stability, we secure the following results, based upon prices during the period 1913-1926 (distributions for the years 1914-1926).

Range of positive values of r	Number of distributions for which the values of r fall within the stated limits	
	Link	Fixed base
7.00, or below	27 out of 52 cases	27 out of 52 cases
6.00, or below	22 out of 52 cases	23 out of 52 cases
5.00, or below	13 out of 52 cases	12 out of 52 cases
4.00, or below	1 out of 52 cases	4 out of 52 cases

The number of link relative distributions falling in the heterotypic area (i. e. having values of r equal to or less than 7.00) is precisely equal to the number of distributions of fixed base relatives in this class. In certain other respects, however, the fixed base relatives are less adequately described by the usual statistical measures (i. e. more fixed base than link distributions fall beyond the boundary marking the upper limit of the validity of the first and second moments). There is, however, slight basis for a choice between them on the basis of this evidence.

Finally, comparing these distributions in respect to conformity to the normal type, we have the following figures. These relate to distributions for the years 1914-1926.

Form of price relatives	Number of distributions deviating from normal type in respect to			Number of distributions conforming to normal type in respect to skewness and kurtosis (within sampling limits indicated in the text)	Total
	skewness alone	kurtosis alone	skewness and kurtosis		
Link	2	7	39	4	52
Fixed base	1	4	46	1	52
Total (including duplicates)	3	11	85	5	104

Four of the five distributions which approximate the normal type are composed of link relatives. The other figures confirm this evidence in indicating a slightly less pronounced departure from the normal type in the case of link relatives than is found for fixed base relatives. But the deviation from the normal law is great for both groups of distributions.

Certain of the results of these comparisons may be summarized.

a. Arithmetic and logarithmic distributions
   1. In respect to dispersion logarithmic distributions are not subject to the extreme fluctuations characteristic of natural distributions during periods of price disturbance.
   2. Logarithmic distributions are more symmetrical than natural distributions.
   3. The tendency toward peakedness is more pronounced among arithmetic than among logarithmic distributions.
   4. Taking all the evidence into account, the logarithmic distributions appear to conform to the stable frequency types (the Gaussian and Type III) somewhat more closely than do the natural distributions. They appear, also, to be capable of somewhat more efficient location and description by the methods commonly employed than are the distributions of price relatives in natural form. The departures from the stable types are, however, very pronounced for both types of distributions.

b. Unweighted and weighted distributions
   1. The dispersion of weighted distributions was generally

smaller during the extreme price disturbance of the war and post-war years. During other periods no consistent difference has been noted.

2. Weighted distributions appear to be more symmetrical than unweighted during periods of price disturbance. At other times no difference is apparent.
3. The use of weights decreases the abnormal peakedness characteristic of distributions of price relatives.
4. Weighted price relatives, in combination, yield more stable distributions than do unweighted relatives. Part of the evidence on this point consists of the fact that 14 out of 95 weighted distributions conformed to the normal type (within sampling limits), while only 4 of 95 unweighted distributions were of the Gaussian type.

c. Distributions of fixed base and link relatives
   1. There is no significant difference between distributions of fixed base and link relatives in respect to symmetry.
   2. Link distributions, for the period 1914 to 1926, are less peaked than are fixed base distributions (the standard of comparison being, in each case, a normal distribution having the same standard deviation as the distribution in question).
   3. There is little difference between link and fixed base distributions in respect to stability of type. The departure from the normal type appears to be slightly less pronounced for link relatives than for fixed base relatives.

In concluding this section the general characteristics of frequency distributions of price relatives may be noted. Classifying these distributions by type, we have:

Type I	19
Type IV	147
Type VI	24

(Of the 19 distributions of Type I, 6 belonged to sub-group I_J, and 12 of the 24 distributions of Type VI belonged to the sub-class VI_J).

One conclusion stands out clearly from a survey of these results. Price relatives are not, in general, distributed in accordance with the normal law of error. This conclusion holds whether we deal with link or fixed base relatives, and whether these relatives be combined in natural or in logarithmic form. Of the entire 190 distributions which have been analyzed, all but 18 differ from the normal type in respect to skewness, or kurtosis, or both, by amounts which are not consistent with the hypothesis that the populations sampled are distributed according to the normal law of error. (See table on p. 331). The sampling errors for which allowance was made in arriving at this figure were large, so that 18 must be looked upon as a generous estimate of the number of distributions which might be described by the normal curve.[1] The great majority of the distributions (136 out or 190) depart from the normal type in respect to both skewness and kurtosis. Of the 36 cases in which one factor alone is significantly non-Gaussian, 30 involve the factor of kurtosis, and only 6 that of skewness. It would appear that the excessive concentration of cases near the modal value is somewhat more important than the factor of skewness in causing distributions of price relatives to depart from the normal type.[2]

The data employed in preparing the distributions analyzed above cover a long period of years and have been combined in a variety of forms. They furnish fairly conclusive evidence concerning the form which frequency distributions of price relatives take. It is clear that the conditions which give rise to the normal distribution are seldom realized in the system of prices, in so far as degree of change in price between specific dates is concerned. The factors affecting such changes in individual commodity prices are not independent and indefinitely great in number. (Or if they are unrestricted in number, it would appear that a limited number are of dominant importance.) Nor does each cause group affecting the items in a given distribution contribute with equal frequency results in excess and in defect of its mean contribution.

The majority of distributions of price relatives are of Type IV of Pearson's classification. Almost one-quarter of the total number

[1]As was noted, this test relates specifically to the hypothesis that the populations sampled are distributed according to the Gaussian law. It does not test whether, from the actual populations, individual samples following the normal law might not be secured. In the case of all distributions represented by points in the heterotypic area there is some possibility of securing individual samples following the normal law because of the infinite probable error of β_2.

[2]In commenting upon the peakedness of distributions of price relatives, to which attention was called by Wesley C. Mitchell, Frederick R. Macaulay has pointed out that this tendency toward heavy concentration is a result of the characteristic inertia of prices (*American Economic Review*, March, 1916, p. 205). This peakedness, which is present in a significant degree in all but 24 of the 190 distributions studied, appears to be a resultant of two factors—the inertia of some prices, of which Dr. Macaulay speaks, and a relatively high degree of variability of certain other prices. The extreme changes of the latter group give the distribution as a whole a degree of variability not consistent (in terms of a normal distribution) with the stability of the former group.

are of other types, however. And of those which are classed as Type IV, many are represented by points in the heterotypic area. Almost half of the total number of distributions (90 out of 190) are, in fact, heterotypic.[1] It is certain that there is no one ideal frequency type to which distributions of price relatives conform.

If we look upon price relatives, irrespective of the years to which they relate, as a single population, they must be classed as of an erratic and extremely unstable type. And if relatives for successive years be regarded as constituting separate populations, these populations appear to differ so radically from each other that they cannot be considered as belonging to a single homogeneous family. This is not to say that comparison of measures relating to different years may not be very significant. It is merely to say that we do not have here a stable, homogeneous population, as such populations are found in handling many types of statistical data in biology, economics and other scientific fields.[2]

Although no pronounced change in frequency type is effected by changing the form of price relatives (i. e., from natural to logarithmic) before combining them, by using weights, or by employing link in preference to fixed base relatives, or vice versa, there is evidence that something is gained in the way of statistical accuracy by such modifications of method. There is some apparent improvement in stability of type and in the efficiency of the orthodox descriptive measures when logarithms of price relatives are used, in place of relatives in natural form, when weights are employed, and when link relatives rather than fixed base relatives are combined.[3]

[1]If a distribution falls in the heterotypic area it means that the fourth moment has an infinite probable error. Truman L. Kelley answers the question as to how a distribution of price relatives could have an infinite feature by pointing out that certain commodities for sale in 1917 were not purchasable at any price in 1918. The population of 1918 price relatives, on the 1917 base, did, therefore, contain infinite ratios. The infinite characteristic of the actual distribution studied as a sample is evidence of the possibility of such infinite ratios (*Statistical Method*, p. 146).

Kelley argues that the instability of distributions departing from the normal type (or from Types III or V) is an instability inherent in the data, not a mere oddity of the equations representing the various distribution types.

[2]Striking evidence of this is found in the extraordinarily high values of β_1 and β_2 which were secured from some of the price distributions. A number of these exceed the highest values of these criteria previously observed. Pearson (*Phil. Trans.*, Vol. 216-A, p. 440) cites as the highest observed values of β_1 and β_2, of which he has heard, those given by Duncker (*Biometrika*, Vol. VIII, p. 238). These are

Armzahl, *Asterina exigua* (N=600) β_1=1.76, β_2=33.13

Armzahl, *Archaster typicus* (N=902) β_1=4.76, β_2=128.48

[3]These statements have to do only with the question of distribution type. Many other considerations, some of which have been touched upon in earlier sections, bear upon the choice of methods in particular cases. It will be observed that in following chronological changes in type in the succeeding section use has been made of unweighted relatives in natural form, as well as of weighted logarithms of relatives.

§On the Form of Frequency Distributions of Price Relatives

There are many statements in the literature on index numbers concerning the type of frequency distribution secured when price relatives are combined. F. Y. Edgeworth, in his *Memoranda* on "Measurement of Change in the Value of Money" (1887-1889) referred to the asymmetrical character of many distributions of price relatives, in natural form, and suggested that the Galton-Macalister curve (the equation to which gives a normal distribution when the logarithms of x-values are taken) might be used to represent such distributions. (See *Papers Relating to Political Economy*, Vol. I, p. 242.) Edgeworth states, however, that this asymmetry is not found in all distributions of price relatives, and that price relatives in natural form would be expected to combine in the form of a normal distribution "where the entries are average prices based on a great number of items."

Wesley C. Mitchell made a detailed comparison of the actual distribution of 5578 link relatives (drawn from different years) and the corresponding normal distribution. (*The Making and Using of Index Numbers*, *Bulletin* 284, *Wholesale Price Series*, U. S. Bureau of Labor Statistics, pp. 18-19.) He finds, in the first place, the asymmetry noted by Edgeworth, and suggests that the use of a logarithmic x-scale would result in a more symmetrical distribution. Even more pronounced, however, was the difference noted by Mitchell between the actual and the normal distributions in respect to peakedness. The actual distribution was much more peaked than the normal. The results of the present study indicate that this leptokurtic distribution is characteristic of price relatives, an even more consistent characteristic than is the positive skewness.

A. W. Flux has suggested the possibility of a study of the changing shapes of distributions of price relatives, and has touched upon the probable form of such distributions. Speaking of twelve-month link relatives, he says "If the quotations are sufficiently independent, the distribution of the 150 variations should, when the general price level is stable, accord with the well-known law of error. If prices are rising the curve showing the 'scatter' of the different price movements should take a skew form with its mode to the right of the mean, and if they are falling the mode should be to the left of the mean." ("The Measurement of Price Changes," *Journal of the Royal Statistical Society*, Vol. 84, 1921, p. 190.) The first condition, that the quotations should be sufficiently independent, is one which would be difficult to ensure in practice. It is possible that the distribution of a selected group of price relatives, picked out on the basis of independence, might accord with the normal law when the price level was stable. In such a selection of prices as is employed in the computation of most index numbers some degree of intercorrelation is to be expected. (Professor Bowley's results have indicated the degree of correlation prevailing in a typical group.) For such a group a close approach to the normal law appears to occur very rarely, even with a stable price level.

The discussion in the succeeding sections bears upon the further suggestions of Mr. Flux concerning the character of the skewness to be expected with rising and falling prices. There appear to be some exceptions to the relations he suggests, since all phases of rising (or of falling) prices are not identical in respect to the nature of the changes in the constituent items.

Lucien March, using the data referred to at an earlier point, tests the conformity of distributions of relatives computed from French, English and American prices to the normal law, He compares the relative frequencies of the actual observations falling within stated limits, measured from the mean, with corresponding frequencies for a normal distribution. (For Professor March's results see *Metron,* Vol. I, No. 4, 1921, pp. 82-3.) He concludes, "Quant à la distribution de ces écarts suivant leur grandeur, on voit........qu'elle se conforme d'une manière relativement satisfaisante à la loi normale." No rigorous test of the significance of the observed differences is applied, presumably because, for Professor March's purposes, a very rough agreement was satisfactory. In the absence of a test more precise than that applied this does not constitute evidence of conformity to the normal law. In comparing the normal and actual frequencies, positive and negative deviations are lumped, and the comparison is solely in terms of the magnitude of the deviations. Since Professor March has shown, in another table, that his distributions are asymmetrical (the percentage of cases above the mean varying from 32 to 58 in the six distributions he studied) it is not valid, for general purposes, to test conformity to the normal law in terms of the magnitude of the deviations alone. Professor March's conclusion, it should be noted, is applied by him only to the magnitude of the deviations.

Corrado Gini ("Quelques considérations au sujet de la construction des nombres indices des prix et des questions analogues," *Metron,* Vol. IV, No. I, 1924) reviews various pieces of evidence on this subject. Though he recognizes the claims of various authorities that price variations follow the normal law, he contends that this conformity is not always found. The present investigation justifies the broader statement that distributions of price relatives very seldom conform to the normal law.[1]

3. Chronological Changes in the Characteristics of Frequency Distributions of Price Relatives

In the preceding section various distributions of price relatives have been compared in respect to those characteristics which define

[1]The book by Dr. Maurice Olivier (*Les Nombres Indices de la Variation des Prix*) to which previous reference has been made contains the results of a comprehensive study of the form of distributions of price relatives in France, between 1920 and 1924. The data are relatives on the 1913 base, computed from the price series entering into the French wholesale price index of the Federal Reserve Board. Dr. Olivier finds that the distributions of these price relatives in natural form are far removed from the normal type. Distributions of logarithms of these relatives are closer to the Gaussian form, but remain distinctly more peaked than the normal curve (pp. 98-154).

population type. From the charts which have been presented it is clear that, for any given type of distribution, these characteristics change materially from year to year. It remains to inquire as to whether these changes possess significance, for the purposes of the present investigation.

Frequency distributions of price relatives may change from year to year in respect to any one, or all, of their characteristics. The central tendency may change—and variations of this type furnish the subject matter of the extensive literature on index numbers of prices. There may be changes in the degree of dispersion, a topic which has been treated in some detail in a preceding section of this volume. The characteristics with which we are now concerned are those relating to population type. These characteristics are defined by the specific measures of skewness and kurtosis and by the several criteria of curve type which have already been described. In the following pages we shall trace the year-to-year changes in these characteristics for two distribution types.

§Changes in the Characteristics of Frequency Distributions of Unweighted Fixed Base Relatives

This discussion may be introduced by a brief statement concerning the interpretation of several of the measures employed. The direction and degree of skewness of a distribution of price relatives for a given year is of considerable importance, particularly when considered in connection with the nature of the change in the level of prices between two dates. The skewness is negative, of course, when the tail of the distribution extends to the left, when low relatives occur more frequently than corresponding high ones. The mean, in such a case, is less than the median in value. With positive skewness the tail extends to the right, high relatives occur more frequently than corresponding low ones, and the mean exceeds the median in value. This measure has significance, for our present purpose, because of the light it throws on the nature of a given change in the level of prices between given dates.

When the skewness is zero with rising prices it means that the advance has been a perfectly symmetrical one, that there has been an upward movement of the bulk of commodities, and that high and low price relatives representing a given deviation from the mean occur with the same frequency. Thus in 1908 the arithmetic average of unweighted relatives of prices, on the 1902 base, was 106.8. The skewness was zero (within sampling limits). We may take this to mean that the advance between 1902 and 1908 had been a perfectly balanced one. The dispersion was considerable, but the distribution of the relatives about the central tendency was symmetrical. A balanced and symmetrical price decline is measured in the same way. Thus in 1894 the arithmetic average of unweighted relatives on the 1891 base was 85.7 and the skewness was zero.

Positive skewness with rising prices means an unbalanced advance, in the sense that a few commodities have risen materially and that the bulk of commodities have lagged behind. This was the situation, in respect to fixed base relatives (1902 base), in 1907. Negative skewness with rising prices means an unbalanced advance of another type, with the bulk of commodities rising, but with a small number falling materially, or lagging behind on the advance. Some will always lag, of course, but the condition mentioned means that the number which lag exceeds the number which lead by corresponding amounts in the price advance. In 1910 such a situation prevailed among relatives on the 1902 base.

Positive skewness with falling prices represents a condition in which the bulk of commodities have moved downward, but a few have risen in price (or have fallen less than the rest). In 1895 price relatives on the 1891 base made up a distribution having this characteristic. Negative skewness develops with falling prices when a few commodities have fallen materially, with no corresponding increases, and with the bulk of articles not falling, or lagging behind the few on the decline. The 1893 distribution of relatives on the 1891 base portrays a condition of this type. The widely different conditions which may accompany a price rise (or a price fall) should be distinguished, if a clear understanding of the price situation is to be obtained.

The measure of kurtosis indicates the shape of a given distribution in another respect. If the concentration at the central tendency is greater than in the corresponding normal distribution, with a smaller concentration at the shoulders of the curve and a wider spread at the tails, this measure is positive in value. If the central concentration is less than in the corresponding normal distribution, the measure of kurtosis is negative. The measure is positive, in general, for distributions of price relatives.

The story of the year-to-year changes in type of the distributions of price relatives is not a simple one for, as has been suggested above, we are dealing with a population which seems to contain inherent elements of instability. This population is subject to violent internal changes which are reflected in radical alterations from year to year in the form and character of the frequency distributions. Yet there is some order in these fluctuations, and information of some value concerning changes in the price system from year to year may be gleaned from a study of these alterations in type.

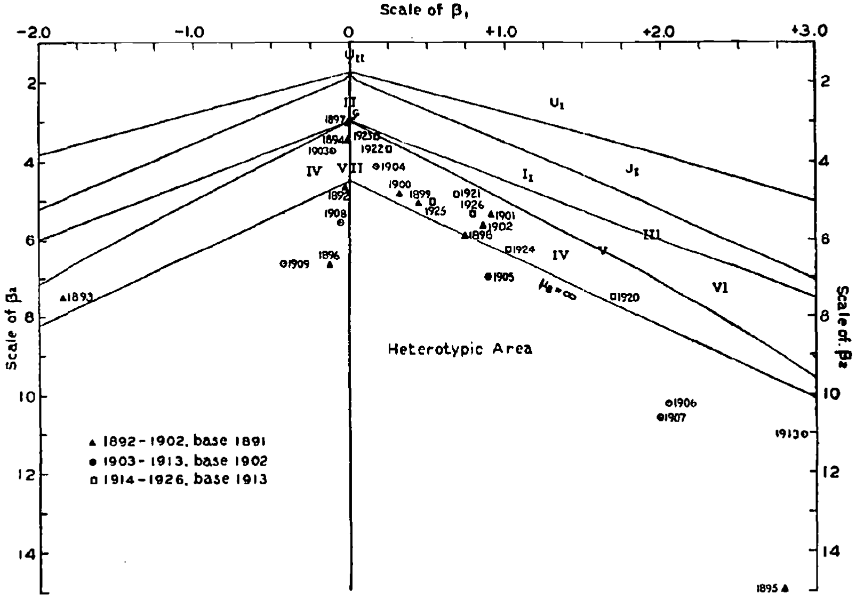
It is desirable to follow in some detail the changes in distributions composed of both fixed base and link relatives. The distributions which have been analyzed for the entire period since 1890 represent combinations of unweighted and weighted fixed base relatives, in natural form and unweighted and weighted link relatives, in logarithmic form. Those which will be discussed here are the unweighted fixed base distributions, and the weighted link distributions.[1]

[1]The year-to-year changes in the criteria of curve type relating to the unweighted natural distributions seem to be more consistent and more significant than the changes in the corresponding weighted measures. Although measures computed from weighted relatives would in general be preferred, certain of the annual changes in criteria of curve

FIGURE 43

THE POPULATION OF PRICES.

Diagram Showing the Location of Points Representing Distributions of Unweighted Fixed Base Relatives in Natural Form, 1892-1926.[1]



[1]The relatives which were combined in forming these distributions fall into the three groups distinguished by the legend.

The points representing distributions for the following years fall beyond the limits of the chart:

Year	β_1	β_2
1910	—5.9280	18.9393
1911	13.6001	31.5601
1912	15.0925	33.2958
1914	8.9019	23.5430
1915	118.4667	139.3550
1916	164.2497	200.3467
1917	147.4448	192.4489
1918	126.6821	148.1861
1919	29.6149	68.2365

Fixed base relatives have been computed for three different periods, so that it is possible to study their behavior, in combination, under

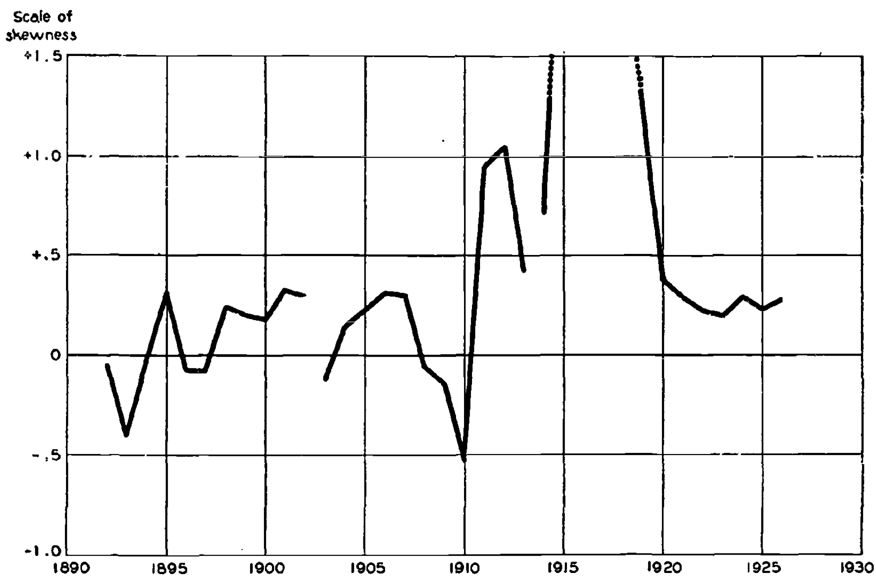
type for this group appear to be more erratic than those of the unweighted relatives. This is probably due in general to wide price fluctuations of one or two heavily weighted articles. Extreme price increases would be given heavy weight, in any case, in arithmetic distributions, and this tendency is intensified when widely varying weights are employed and when the higher moments are involved in the calculations. The use of logarithms gives less weight to high relatives, and permits advantage to be taken of weights without introducing the erratic movements that may be present when natural numbers are weighted. The differences between weighted and unweighted distributions are not extreme, however.

The statement that the year-to-year changes in type of weighted natural distributions may be erratic does not mean that they may be extreme (in comparison with unweighted distributions). They appear at times to be accidental, inconsistent with previous movements, and are apparently the reflection of extreme movements in the prices of a few commodities. For this reason they may not have great economic significance markedly different conditions. The first part of the first period, 1891-1896, was marked by falling prices and generally depressed industrial conditions. This was followed by recovery and a pronounced price rise during the later years. The second period, 1902-1913, was characterized by a sustained rise in the price level, broken by two short periods of price decline. The third period, 1913-1926, covers the price revolution of the war and post-war years.

The changes in the character of distributions of unweighted fixed base relatives may be followed on Figure 43, which shows the location by years, of the points defined by the values of β_1 and β_2. Since the chronological changes are difficult to follow on these charts, the information which they yield is presented, in a slightly different form, in Figures 44 to 46. Departures from the normal type due to the lack of symmetry[1] of the successive distributions are shown in Figure 44. The movements of the measure of skewness above and below the zero line trace the year-to-year fluctuations of the points plotted in Figure 43 as they swing to the right and left of the axis of symmetry. The next chart (Figure 45) traces variations in the degree of peakedness[2] of the dis-

FIGURE 44

MEASURES OF SKEWNESS OF DISTRIBUTIONS OF UNWEIGHTED FIXED BASE RELATIVES IN NATURAL FORM, 1892-1926.



[1]Pearson's measure of skewness $\dfrac{\sqrt{\beta_1}\,(\beta_2 + 3)}{2\,(5\beta_2 - 6\beta_1 - 9)}$ has been employed.

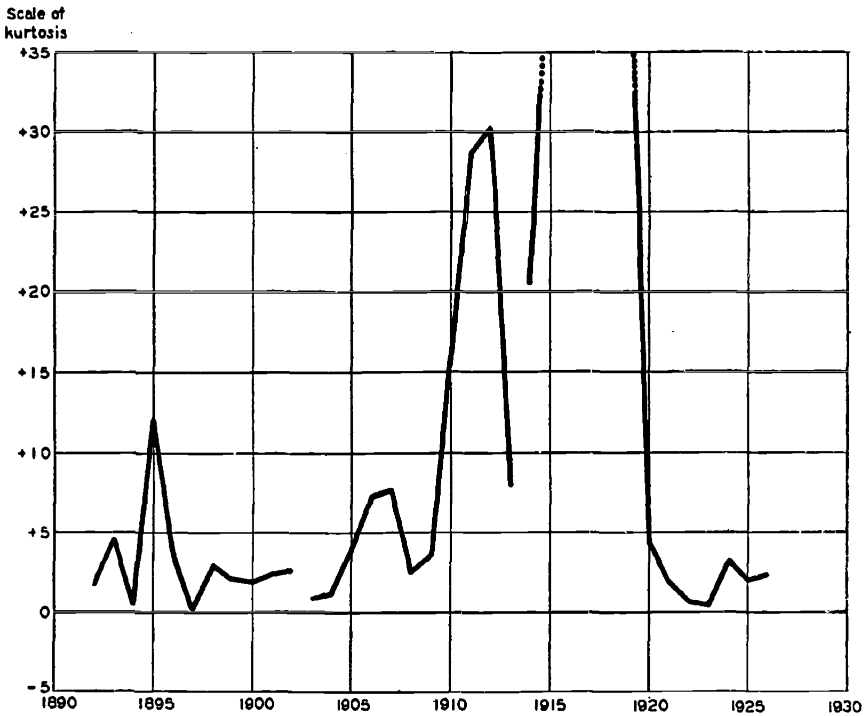
[2]As measured by $\beta_2 - 3$.

tributions of fixed base relatives in natural form. These two figures (44 and 45) thus reflect deviations from the normal type in the two basic respects in which distributions may vary from the Gaussian form. The criterion κ_1, which measures degree of departure from Type III, is plotted in Figure 46. Type III, it will be recalled, includes those stable frequency distributions for which all positive moments are finite. In tracing the fluctuations from year to year of the several measures just discussed, concurrent changes in the central tendency and in the dispersion may be noted. These measures are given in Appendix Table XIX.

Following in Figure 43 the movements of the Beta points representing the various annual distributions, and tracing the year-to-year changes in skewness and kurtosis, we have a story of recurrent swings away from and back to a balanced condition of symmetry and stability.

FIGURE 45

MEASURES OF KURTOSIS OF DISTRIBUTIONS OF UNWEIGHTED FIXED BASE RELATIVES IN NATURAL FORM, 1892-1926.

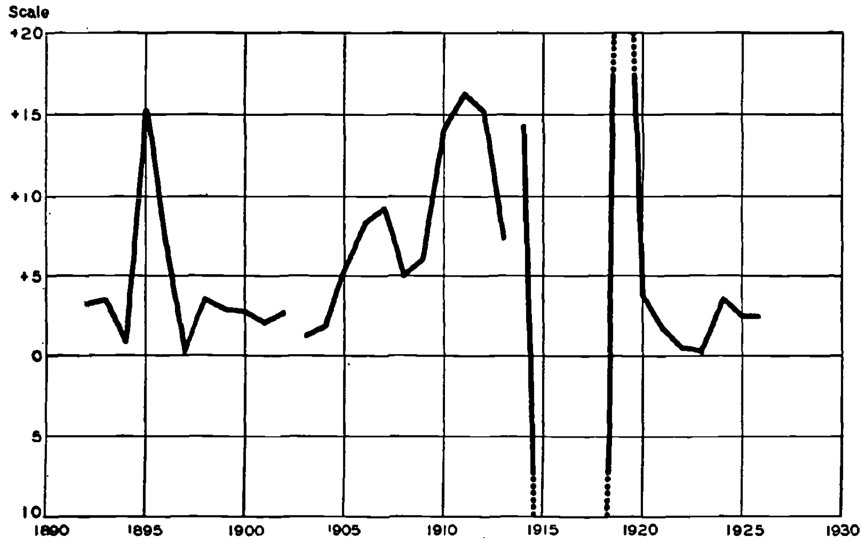


In the first period (1891-1902) there are two such swings, and in the second period (1902-1913) there are also two. During the period from 1913 to 1926 there is one broad swing. The movements correspond in general to the major cycles in business.

The observations for the first period begin with the distribution for 1892. From 1891 to 1892 there had been a balanced decline in average prices. Although there was no material change in the price level between 1892 and 1893 there was a relatively severe internal disturbance. The distribution in 1893 of 1891 base relatives is sharply skewed in a negative direction, and the Beta point is pulled far away from the Gaussian point. It will be recalled that 1893 was a year of panic and depression.

FIGURE 46

VALUES OF THE CRITERION κ_1 FOR DISTRIBUTIONS OF UNWEIGHTED FIXED BASE RELATIVES IN NATURAL FORM, 1892-1926.



Between 1893 and 1894 there was a material price decline, but it was one which brought a return to symmetry in the form of the frequency distribution and a close approach to stability. The distribution of fixed base relatives in 1894 is, within the limits of random sampling, of the Gaussian type. By 1894, it appears, the effects of the sharp disruptions of 1893 had been repaired. In 1895 there came another violent disturbance, but this time the skewness is positive. The Beta point in Figure 43 is carried far to the right of the axis of symmetry and into the heterotypic area. (1895 was a year of sharp price rise, and a sharper fall. The net change in the price level, as compared with 1894, was negligible, but the characteristics of the frequency distribution underwent a great change.) Symmetry and balance were restored in 1896 and 1897, the years which marked the lowest point of the price drop of the 1890's. The distribution for 1897 is one of the few definitely Gaussian distributions found. The general price advance which began after that date brought five successive distributions, (for the years 1898 to 1902) which are clustered in a small area slightly to the right of the axis of symmetry. All are somewhat skewed in a positive direction, but there are no extreme irregularities in these distributions. The year 1900, during which a minor recession set in, is closest to the axis of symmetry.

The second period, which covers the years 1902-1913 (the relatives being on the 1902 base), witnesses a similar series of swings away from the axis of symmetry, the swings culminating in 1906-07 and in 1911-12. The period starts with a negatively skewed distribution in 1903. For the four years following the skewness is positive, reaching its maximum in 1906. The changes in respect to symmetry were accompanied by increases in the degree of peakedness of the distributions, and by a constantly greater departure from the Gaussian point. That is, during this four year period there was an unbalanced upward movement of prices, with a few extreme advances extending the tails of the distributions to the right.[1] The distributions of 1906 and 1907 are clearly heterotypic. With the depression of 1908 came a swing back to symmetry. There is a distribution of slight negative skewness for 1909, and 1910 is more sharply skewed in the same direction. Following that there was another sustained swing to the right, with a succession of asymmetrical distributions, deep in the heterotypic area. Not until 1913 was there another movement in the direction of symmetry, and this did not carry all the way back. In these years (1911, 1912, 1913) the bulk of the commodities lagged behind the few in the advances that had been made since the base year (1902).

During the third period (1913-1926, 1913 being the base) there was a single major swing away from and back to the axis of symmetry. The distributions of fixed base relatives are positively skewed in all years. This means that for this entire period there is a preponderance of frequencies above the central tendency (in the sense that extremely high relatives occur more frequently than corresponding low ones). The degree of asymmetry varies materially, however, for different years, and the degree of departure from the Gaussian point varies even more widely. The swing away from symmetry and stability began in 1914, reaching a maximum in 1916. The chief characteristic of the distributions for 1915 and 1916 is the extreme elongation of the tails extending to the right, which means that the bulk of commodities lagged far behind on the rise. During the years 1917, 1918, and 1919 a tendency to correct this lack of proportion is in evidence. The bulk of the commodities moved up, and the most extreme departures were reduced in magnitude. The distributions remain skewed and unbalanced, however, to a degree never approached during the other years of the period covered.

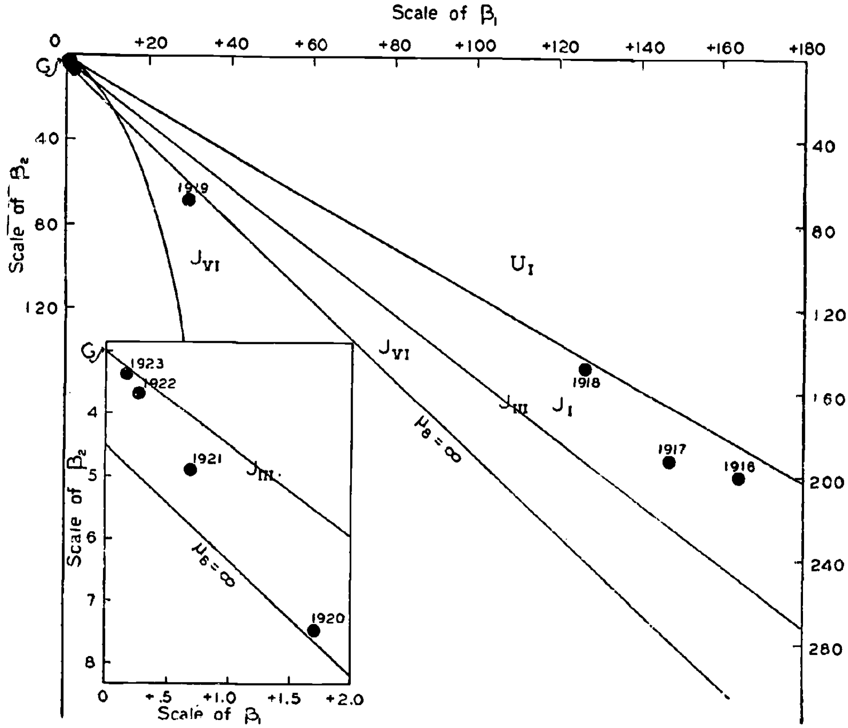
The changes in the character of distributions of price relatives which occurred between 1916 and 1923 are of exceptional interest, and for this

[1]A distribution may depart widely from the Gaussian type because of the influence of one or two extreme price changes. This is notably true of the distribution relating to the year 1915. One should be cautious in basing conclusions upon changes in type due to the influence of a small minority of the observations, but it is a fact of great significance that price distributions are materially affected by a few divergent observations, and that such divergencies may occur among apparently homogeneous data.

FIGURE 47

THE MOVEMENT TOWARD PRICE EQUILIBRIUM, 1916-1923.

Diagram Showing the Location of Points Representing Distributions of Unweighted Fixed Base Relatives in Natural Form[1].



The insert in the lower left hand corner, which is an enlargement of that portion of the main diagram in the neighborhood of the Gaussian point, shows the changes between 1920 and 1923.

[1]Base of relatives: 1913.

reason the points representing these distributions are shown in Figure 47, on scales which permit the whole movement to be followed. By 1916, as has been noted, the swing away from the axis of symmetry had attained its peak. During the seven years between 1916 and 1923 there was an unbroken movement back toward the normal type of distribution. In 1917 and 1918 this movement was not pronounced; 1919 brought an acceleration of this return, though the distribution for that year is still sharply skewed in a positive direction. The year 1920 marks an equally pronounced regression toward the normal type. The points representing the distributions for 1920, 1921, 1922 and 1923 appear in the insert in Figure 47, in which the scales have been enlarged. These four points lie practically on a radial line from the Gaussian point, and come progressively closer to it. A normal curve could be employed to describe the distribution for 1923. (The distribution in 1923 of weighted relatives on the 1913 base gives an almost perfect Gaussian distribution.) The dispersion is wide, as may be seen by reference to the frequency distribution itself (Figure 21), but it gives evidence of being the resultant of a balanced play of forces. The distribution for 1922 is very close in type to the 1923 distribution. By 1922, we may say, the distortion and the disturbed conditions which the price revolution of the war and immediate post-war years had brought had been in large part corrected. The various forces which were acting upon the 1913 base relatives had by that year become so numerous and so well-balanced that their interaction was able to bring about that orderly distribution to which the name of Gauss has been attached.

It is a curious and dramatic thing, this march of the successive Beta points back to the Gaussian point, as the conditions of balance and independence and multiplicity of causal factors which are characteristic of the normal law, and which had been so markedly absent during the war years, were gradually re-attained in the price system. It exemplifies in striking fashion the emergence out of chaos of that form of order which the play of sheer chance brings. Here is nature forming habits.

The years 1924 to 1926 brought a minor swing away from the axis of symmetry in the direction of positive skewness. The points representing distributions for these years are within or close to the little circle within which lie the points for the five years of slowly rising prices and of comparative quiet between 1898 and 1902. The later distributions are of the same basic type as these earlier ones.

The conclusions concerning this tendency of distributions of fixed base relatives to recover from the disturbances attendant upon extreme price changes have been checked by a study of distributions of unweighted relatives on the 1891 base (in natural form) carried forward through 1926. The number of commodities included was 195 for the entire period (except for 1918 and 1925, in which prices for only 194 were available, and 1926, for which 193 were included). The descriptive measures are given in Appendix Table XXVII.[1]

During the decade prior to the war the distributions of relatives on the 1891 base differ in the details of their movements from the distributions of relatives on the 1902 base. The positive skewness is somewhat more pronounced, and the recession of 1908 does not bring the return to symmetry which it did for the distributions on the later base. Not until 1909 did the points representing distributions on the 1891 base swing back toward the axis of symmetry. During the war and post-war years, however, we find the distributions of 1891 base relatives following much the same course as the distributions on the 1913 base, although the disturbances of the war years had less marked effects on the relatives resting on the distant base. The price changes of 1915 to 1918 brought the same series of positively skewed distributions, and the years 1920 to

[1]The fact that the number of observations upon which these distributions are based is smaller for the period since 1913 than the number included in the distributions on the 1913 base renders the criteria of curve type somewhat less reliable, i. e. subject to greater sampling fluctuations.

1923 witnessed the same pronounced movement back toward symmetry and stability. In this case the movement did not carry as far back toward the normal type. It is noteworthy, however, that the distributions for 1921, 1922, 1923 and 1924 are, within the limits of sampling, of Type III. This type may be considered to stand next to the Gaussian in respect to stability. It is the form of distribution which arises when, in Pearson's terms, "each cause group is of equal valency and independent, but does not give contributions in excess and defect (of the mean) of equal frequency." It appears from this evidence that the factors which from 1921 to 1924 shaped the distribution of relatives on the 1891 base were numerous (or, if limited in number, were of equal importance) and independent. The conditions affecting these distributions differed from those which bring about a normal distribution in that the contributions of individual factors above and below their mean contribution did not tend to be of equal frequency. One important result of the price changes during the thirty years which had elapsed since the base year was the excess of contributions above the mean, giving the positive skewness which has been noted. But, in spite of this, the elements of instability (represented by infinite positive moments) which appear in most distributions of price relatives are not present in these four distributions. (Type III distributions, to which these approximate, are characterized by finite positive moments of all degree.) It may be hazarded that this curious re-appearance of stability of type after thirty years, during which violent price disturbances took place, is due to an increase in the number of factors affecting the prices of the commodities studied, a corresponding decline in the importance of a limited number of factors which may at times have exerted a preponderant influence, and a decline in the strength of the varied intercorrelations between factors. The mere passage of time might be expected to bring about just such changes as these, and the evidence furnished by the various fixed base distributions which have been analyzed indicates that such a tendency is clearly present.

But it is not a permanent condition of stability which is thus brought about. Just as the distributions of relatives on the 1913 base swing away from the normal type after 1923, so the distributions of 1891 base relatives depart from the type which had prevailed during the four years from 1921 to 1924. The business and price developments of 1925 and 1926 introduced new factors, and distribution types deviating widely from the Gaussian and from Type III again appeared. But, if the history of past changes is repeated, time will bring further corrections, disruptive elements will be eliminated, or balanced, and there will be a new movement in the direction of the stable frequency types.

§ Changes in the Characteristics of Frequency Distributions of Weighted Link Relatives

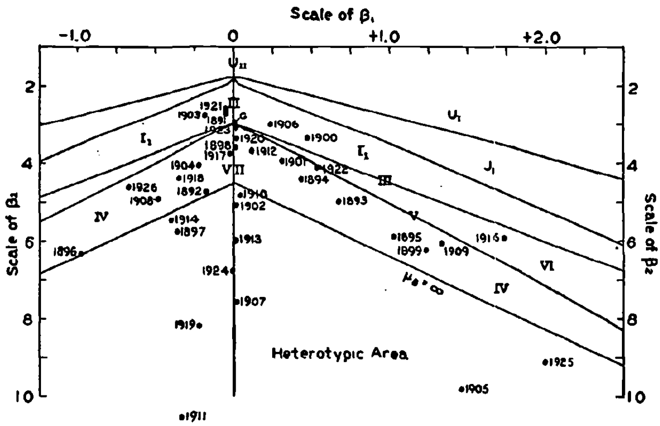
Figure 48, shows the changes in the location of the points representing distributions of weighted logarithms of link relatives of wholesale prices, as these points are defined by successive values of β_1 and β_2.

Changes in skewness, kurtosis and stability (as measured by κ_1) may be followed on Figures 49 to 51.

FIGURE 48

THE POPULATION OF PRICES.

Diagram Showing the Location of Points Representing Distributions of Weighted Link Relatives in Logarithmic Form, 1891-1926[1].



[1]The point representing the distribution for 1915 falls beyond the limits of the chart. The values of the criteria are:

$$\beta_1 = 3.7787$$
$$\beta_2 = 20.2149$$

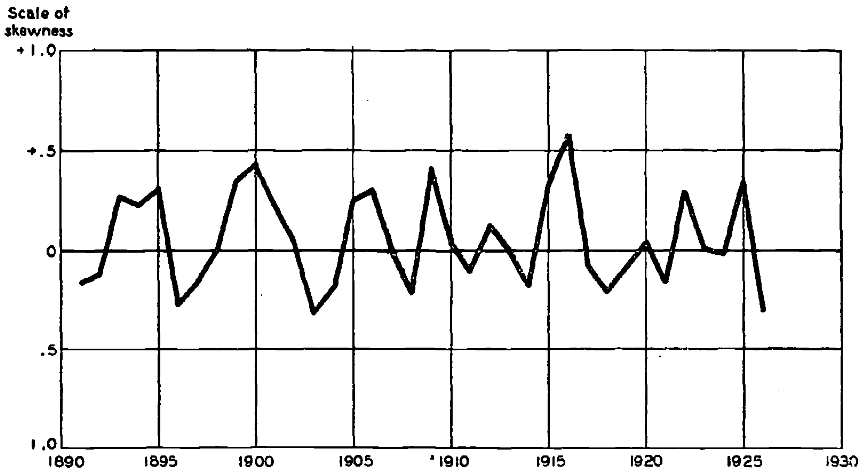
Variations from year to year in the characteristics of distributions of link relatives are, as might be expected, somewhat more pronounced than the variations in the form of distributions of fixed base relatives. Nevertheless, a series of cyclical swings away from and back to the axis of symmetry is discernible in these fluctuations. These swings may be followed most readily in the varying values of the measure of skewness (Figure 49). Tracing these values, for distributions of weighted logarithms of link relatives, we find a series of peaks in the direction of positive skewness in the years 1893-95, 1899-1900, 1905-06, 1909, 1912, 1915-16, 1922, and 1925. The valleys (i. e. conditions of negative skewness) come in the years 1891, 1896, 1897, 1903-04, 1908, 1911, 1914, 1918, 1921 and 1926. The years of approximate symmetry are 1898, 1902, 1907, 1910, 1913, 1917, 1920, 1923, 1924.

If we start with the depression of 1896 there is observable a fairly consistent relationship between the direction and degree of skewness on the one hand and the changes in business activity on the other. The years of pronounced positive skewness are, in general, the years when the tide of business was running steadily in the direction of prosperity, with no material movement in the other direction. This is true of the years

1899, 1905 and 1906, 1909, 1912, 1915-16, 1922 and 1925. (The present evidence places 1900 in this group, although the annals of business show a mild recession in this year.) These are the years when the Beta points were pulled farthest to the right, away from the axis of symmetry[1].

FIGURE 49

MEASURES OF SKEWNESS OF DISTRIBUTIONS OF WEIGHTED LOGARITHMS OF LINK RELATIVES, 1891-1926.



Distributions which are sharply skewed in a negative direction (i. e. with tails extending to the left) are found, in general, in years in which the tide of business activity was receding, or at its lowest point preceding revival. Such conditions prevailed in 1896 and 1897, in 1908, 1911, 1914 and 1921, years in which the Beta points representing weighted logarithms of link relatives in Figure 48 are pulled to the left of the axis of symmetry. In 1903, 1904 and 1918 distributions of the same type are found, and there is evidence in the records of business of similar tendencies in these years, although the tendencies were not pronounced. The negatively skewed distribution for 1926 reflects the general decline in prices during that year, a decline which was exceptional in that it was not coincident with any widespread recession in business.

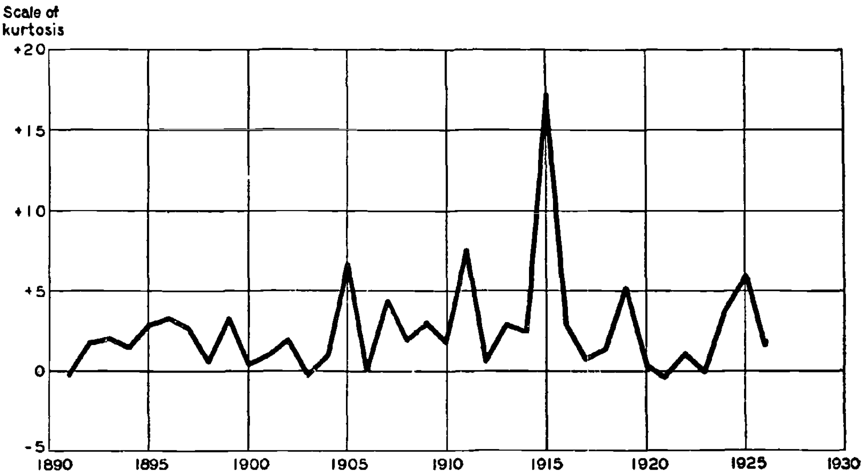
When there are cross-currents in the business world or, more exactly, when the direction of movement changes in the course of the year, symmetrical distributions are secured. This is true of the distributions for 1902 (price index turned downward after October), 1907

[1]Positive skewness means, in any case, that deviations above the central tendency exceed in number corresponding deviations below. When the distributions are logarithmic, the word "corresponding" is to be interpreted in terms of ratios (i. e. a deviation of 20 per cent below a given value is equivalent to a deviation of 25 per cent above). A given degree of positive skewness on the logarithmic scale represents, thus, a greater degree of skewness on the natural scale.

(price index turned downward after October), 1910 (price index turned downward after April), 1913 (price index turned downward after September), 1920 (price index turned downward after May), 1923 (price index turned downward after April), and 1924 (price index turned upward after June). All of these years except the last, it may be noted, are years during which prices turned downward. It is apparent that a downturn within a calendar year tends to produce that condition of balance among price relatives which is lacking when the tide of price change is running steadily up or down.[1]

FIGURE 50

MEASURES OF KURTOSIS OF DISTRIBUTIONS OF WEIGHTED LOGARITHMS OF LINK RELATIVES, 1891-1926.



The dates mentioned above include all but two of those during which symmetrical distributions were obtained. The exceptions are 1898 and 1917. In each of these years the general price index was rising, but there were conflicting forces at work. In 1898 the effects of the preceding depression were still felt, and full-fledged prosperity had not yet arrived. In 1917 the flood of war-time advances was checked by federal price regulation, a factor which served, apparently, to produce the balance which ordinarily results from a turn in prices during the year.

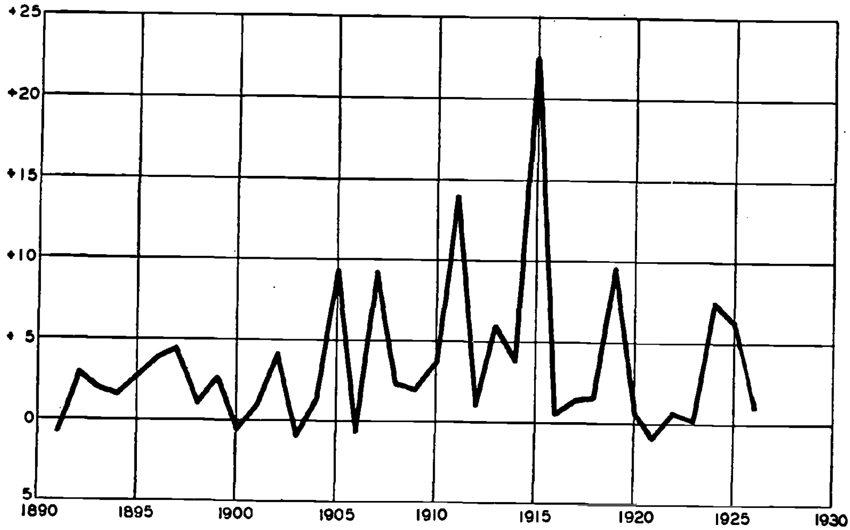
In this discussion the years prior to 1896 have been omitted. The distributions of weighted logarithms of link relatives were negatively skewed in 1891 and 1892, and positively skewed in 1893, 1894 and 1895. Changes in distribution type during this disturbed five year period are not consistent with those noted for the later years, nor is there agree-

[1]The slight decline toward the end of 1925 was not sufficient to offset the pronounced positive skewness of the distribution of that year.

ment between the changes in form of weighted and unweighted distributions[1].

FIGURE 51

VALUES OF THE CRITERION κ_1 FOR DISTRIBUTIONS OF WEIGHTED LOGARITHMS OF LINK RELATIVES, 1891-1926.



[1]The preceding analysis has been restricted to distributions based upon average annual prices. It would be possible to obtain a more accurate account of the changes which frequency distributions of price relatives undergo if we used monthly instead of annual prices, but the labor involved is prohibitive. Monthly price relatives of any of three different types—fixed base relatives, link relatives on the month preceding as base, or link relatives on the twelfth month preceding as base—might be employed in such a study. Distributions of fixed base relatives, it may be expected, would have somewhat the same characteristics as distributions of annual averages, expressed as relatives on a fixed base, but changes in distribution type could be followed with more accuracy. Distributions of month-to-month link relatives would undoubtedly be very peaked (i. e. marked by heavy concentration at the central tendency), somewhat erratic in their changes and, it may be hazarded, would not be of great value in interpreting economic changes. Distributions of link relatives on the twelfth month preceding as base might well prove to be quite significant, if their changes were analyzed in detail. (The utility of measures of changes in the price level, in dispersion, and in displacement computed on this basis has been suggested in preceding sections.) Such distributions would doubtless resemble distributions of annual link relatives in their general characteristics, but should be more sensitive and more accurate indexes of economic change.

Some indication of the form which such distributions would take is furnished by the following figures, relating to distributions of monthly link relatives and twelve-month link relatives. Each of these distributions is composed of unweighted logarithms of relative prices.

Distribution of	N	Mean	Index of dispersion	β_1	β_2	Skewness	Kurtosis
Monthly link relatives (Dec. 1924, on Nov. 1924, as base)	388	102.4	3.5	1.3451	10.7808	+.218	7.78
Twelve-month link relatives (Dec. 1924, on Dec. 1923, as base)	389	103.7	13.2	.4554	3.9329	+.289	.93

The story of the chronological changes in the form of distributions of price relatives has been confined to two particular types of distributions (unweighted fixed base relatives in natural form, and weighted logarithms of link relatives). The following summary contains the chief conclusions to which a study of these changes leads.

1. Year-to-year changes in the form of distributions of fixed base relatives trace a series of swings away from and back to the axis of symmetry. Sharp rises and declines in prices, such as those which occurred in 1893, 1895, 1916 and 1917, distort these distributions and carry them far from the Gaussian type. These deviations may be in the direction of positive or negative skewness. With very rare exceptions they bring distributions more peaked than the normal. Swings away from the axis of symmetry are followed, in the history of price changes since 1890, by clear movements back toward a stable, symmetrical type of distribution, approaching in several instances the Gaussian type. Such returns to the normal (or to a symmetrical) type may come with falling prices, as in 1894 or 1897, or with rising prices, as in 1923 (i. e. falling or rising with reference to the base period). They appear to come when approximate stability has been attained, after a period of disturbance. The tendency of fixed base relatives to cluster in this relatively stable form after periods of violent change, during which extremely skewed and unstable distributions were formed, throws light on the nature of the forces affecting the system of prices, and on the character of their interaction. There is evidence here that after periods of disturbance in the price system there is a tendency toward a gradual restoration of the conditions of independence and multiplicity of causal factors which are necessary to the fulfillment of the normal law, or of the law which is realized in distributions of Type III. The stable types do not persist, year after year, but tend to recur after unstable types have been generated in times of disturbance.

   The clearly defined regression to the normal type between 1916 and 1923 (shown graphically in Figure 47) furnishes a striking illustration of the re-attainment of equilibrium in the price system after a violent disturbance.

2. Somewhat similar swings about the axis of symmetry are observable in the year-to-year changes in the distributions of link relatives in logarithmic form, but a rather different inter-

pretation of these movements is called for. In general, departures from symmetry in the direction of positive skewness are found in years during which business activity is steadily increasing, with few conflicting movements. Swings to the left of the axis of symmetry come in years of recession or of depression, years during which the forces of business reaction are in the ascendant. Symmetrical distributions appear to be characteristic of years during which there are reversals of business and price trends.

3. Distributions of the normal type are somewhat exceptional, but approximations to such distributions are found under conditions which may be defined with some degree of precision. These conditions are not the same for fixed base and for link relatives. Distributions of fixed base relatives in natural form appear to approach the normal type periodically, these periodic returns following intervals of sharp disturbance. The forces which tend to form symmetrical distributions, which at times are of the Gaussian type, re-assert themselves after such disturbances. Among link relatives in logarithmic form that state of balance which leads to symmetrical distributions (and, in perfection, to the Gaussian distribution) is not the typical condition, but seems to arise during years marked by conflicting tendencies. When this balance between opposing forces is attained it is as a transitory condition, appearing to mark a passage from one type of distribution to another. The periodic recurrence of those distribution types which appear under conditions of equilibrium is of particular interest as illustrating a form of moving economic equilibrium.[1]

4. A distinguishing feature of distributions of price relatives is that they do not conform to any one type, but undergo marked changes with variations in price and business conditions. Most distributions of homogeneous data from biological, anthropological or other scientific fields, tend to conform to a common type, without sharp variations from sample to sample. Distributions of price relatives may be of the stable Gaussian type at one time; they may, at another date, be sharply peaked and badly skewed in a positive direction, lying deep in the hetero-

[1]This equilibrium, which relates only to commodity prices and to relative changes in these prices, is not, of course, to be confused with the type of moving equilibrium dealt with by Professor Henry L. Moore, in his recent notable extension of the general theory of economic equilibrium ("A Theory of Economic Oscillations," *Quarterly Journal of Economics*, November, 1926).

typic area. Under other conditions we may find them equally distorted in the direction of negative skewness. The liability of such data to material changes from year to year in their group attributes is a fact of considerable economic importance.

Since frequency distributions of price relatives are subject to marked changes in type, the problem of sampling is a particularly difficult one. During years in which the universe of price relatives is distributed in accordance with the normal law, or years in which their distribution follows any fairly stable frequency type, successive samples may be expected to possess common group attributes, and to yield statistical constants differing but slightly in value from sample to sample. In other years, when the sample distributions fall within the heterotypic area, the evidence suggests that the universe of price relatives is highly unstable. In such cases there can be no assurance that successive samples would possess common group attributes, or that statistical constants derived from the higher moments would approach each other closely in value, when computed from different samples.[1]

VII Relations Among Measures of Price Instability

The various quantities described in the preceding pages have been presented as measures of different kinds of price instability. It remains to determine whether there are any consistent relationships among these measures. Some attention has been given in an earlier section to one phase of this question, the relation between changes in the price level and variations in the degree of dispersion of price relatives. Our present problem is the broader one of measuring relationships among all the measures relating to price stability, whether it be stability of the price level, stability of internal relations, or stability of distributions of price relatives in combination.

In the following tables certain of the results previously discussed are summarized, with additional measures relating to other

[1]The statistical constants commonly employed in studying price movements (the mean and the standard deviation) involve only the first and second moments. These are stable over a wider range than are the higher moments from which the criteria of curve type are derived, and are not subject to the same wide sampling fluctuations. Note should be made, however, of the limited validity of the first two moments for distributions of the J-type which have been found to occur occasionally among price relatives. (Cf. R. A. Fisher "On the Mathematical Foundation of Theoretical Statistics," *Phil. Trans. of the Royal Society of London*, Vol. 222, pp. 338-355.)