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: The Description of Price Relatives in Combination

Chapter Author: Frederick C. Mills

Chapter URL: http://www.nber.org/chapters/c5328

Chapter pages in book: (p. 219 - 229)

those other types of price instability described above—instability of internal relations, and group instability. Measures of the dispersion and displacement of price relatives must be constructed, and account must be taken of those long-term shifts which are due to differences in trend. In the study of group stability and homogeneity we shall be concerned with the attributes of frequency distributions of price relatives. In approaching this problem we lose sight of individual commodities and our sole interest attaches to the attributes of the entire population, considered as an entity.

Somewhat similar problems present themselves when we pass from price relatives to other measures of price behavior. We shall study cyclical measures and measures of variability and trend in combination, again seeking information concerning the group attributes of prices. It would be desirable to analyze in the same way collections of measures of price flexibility and of regional price differences, but the available figures are too limited in number to permit of generalizations concerning group behavior. The study of such measures in combination must wait upon the assembling of the necessary data.

The final stage of the study, and the most important from the point of view of one interested in the elements and internal structure of the price system, is the isolation and analysis of price groups and the testing of various grouping principles. The data described in the present volume furnish some of the basic materials for such an analysis. The detailed account of this part of the study is deferred to a second volume.

II The Description of Price Relatives in Combination

The present chapter deals with the first of the measurable characteristics of commodity prices which were discussed at an earlier point—degree of change in price between specific dates. The first step in the study, in combination, of the relatives which measure these changes is their organization in the form of frequency distributions.

1. Frequency Distributions of Price Relatives

Distributions of price relatives may be constructed in a number of different forms, representing various combinations among measures of the following types:

(a). Fixed base and link relatives
(b). Unweighted and weighted relatives
(c). Relatives in natural and in logarithmic form

It will be desirable to compare the results secured from the employment of certain of these alternative methods.

To facilitate the study of distributions of fixed base relatives the interval covered has been broken up into three periods: 1891-1902, with 1891 as base; 1902-1913, with 1902 as base; and 1913-1926, with 1913 as base. In addition, the prices of a number of commodities during the years 1903-1926 have been reduced to relatives on the 1891 base. In dealing with link relatives, prices in all years from 1890 to 1926 have been employed.

Complete descriptive measures relating to the following distributions of price relatives appear in Appendix Tables XIX to XXVII:

Fixed base relatives, natural form, unweighted and weighted, by periods, 1892-1926

Fixed base relatives natural form, unweighted, 1903-1926 (A special study of 195 price series, reduced to relatives on the 1891 base. Measures for this group for the years 1892-1902 are included under the preceding item.)

Fixed base relatives, logarithmic form, unweighted and weighted, 1914-1926

Link relatives, natural form, unweighted and weighted, 1914-1926

Link relatives, logarithmic form, unweighted and weighted, 1891-1926

Measures of central tendency, dispersion, skewness and kurtosis, and the Pearsonian criteria of curve type have been computed for the above distributions.[1] In addition, measures of central tendency and dispersion have been computed for fixed base relatives in

[1]The construction of logarithmic frequency distributions and the computation of the corresponding logarithmic measures would be a very laborious procedure if the logarithms of the individual price relatives had to be looked up before the measures were tabulated. This may be avoided, and all the ease of computation possible with frequency distributions of the arithmetic type may be secured, by constructing a tabulation form in which the class-limits are stated in natural numbers, but in which the class-interval is constant on a logarithmic scale. For example, if it is decided, upon inspection, that a logarithmic class-interval of .03 would be appropriate for the classification of the logarithms of a given set of relative prices, a form may be prepared upon which the class-limits in logarithmic form appear in the first column. In the second column may be written the natural numbers corresponding to the logarithmic class-limits. The actual tabulation may now be carried out, using the price relatives in their natural form. In the calculations of means and standard deviations, and of the higher moments, the class-interval unit may be used throughout in the customary fashion.

This method is of considerable practical importance, for by its employment geometric means and all corresponding logarithmic measures may be computed as readily as the arithmetic measures. Such logarithmic measures of central tendency and dispersion are possibly of wider general utility than the arithmetic measures usually employed.

logarithmic form, unweighted and weighted, from 1892 to 1913, and for link relatives in natural form, unweighted and weighted, from 1891 to 1913. For the measurement of changes in the price level and for the study of internal instability we thus have eight sets of measures for the years 1891 to 1926 (the initial year is 1892, in the case of fixed base relatives). The number of price relatives employed was 195 for the years prior to 1902, 205 for the years from 1902 to 1913, and from 385 to 391 for the years from 1914 to 1926. The commodities included are described in detail in Appendix Table I.

2. Weights

The weights which have been employed in the present study are based upon the estimated values of the various commodities entering into trade during the period 1920-23, as given in the wholesale price bulletins of the United States Bureau of Labor Statistics. In certain cases the weight of a given commodity has been altered somewhat for the different years covered, in order that the group in which it falls might be adequately represented. For example, raw cotton is represented by two quotations during the years since 1913 and by but one quotation during the years preceding. The one continuing quotation is given heavier weight during the early than in the later period.

It is recognized that weights based upon post-war values in exchange are likely to be somewhat in error, if applied to price quotations extending back to 1890. Accordingly, the weights employed in the present study are to be looked upon only as approximations to ideally perfect weights. No attempt has been made to secure perfect accuracy for each period in the matter of weighting. This would not have been possible, in any case, nor was it considered necessary for the purposes in mind. It has been well established that refinements of precision in weighting are not necessary to the accuracy of price index numbers.

The weight given to each commodity in each period is shown in Table I in the Appendix.[1] The percentage distributions of these weights among the different groups represented in the Bureau of

[1]Only the commodities for which weights are given in Table I have been used in the study of frequency distributions of price relatives. The other commodities listed in this table have been employed in other parts of the present investigation, in which the problem of weighting did not arise.

Weights are applied in all cases directly to the price relatives. Thus if a commodity has a weight of 10, its price relative in a given year is handled as though it occurred 10 times.

Labor Statistics and Federal Reserve Board classifications[1] appear in the following tables. For purposes of comparison the percentage distributions of the weights used by the Bureau of Labor Statistics in 1909, 1919 and 1924, and by the Federal Reserve Board in 1913 and 1923, are also shown.

TABLE 85

PERCENTAGE DISTRIBUTIONS OF WEIGHTS AMONG COMMODITY GROUPS
UNITED STATES BUREAU OF LABOR STATISTICS CLASSIFICATION

(1) Commodity group	(2) Percentage of total weight Present study	(3)	(4)	(5) Percentage of total weight U. S. B. of L. S. index*	(6)	(7)
	1890–1902	1902–1913	1913–1926	1909	1919	1924
Farm products	31.7	31.0	27.1	27.6	28.1	28.5
Foods	22.9	22.4	23.0	26.3	24.5	23.0
Cloths and clothing	11.6	11.7	10.4	11.2	14.3	9.8
Fuel and lighting	13.7	13.4	12.8	10.3	8.7	13.6
Metals and metal products	5.5	6.5	8.5	5.7	6.2	8.4
Building materials	5.4	5.9	5.7	11.5	9.7	5.4
Chemicals and drugs	1.2	1.2	2.0	1.3	1.1	1.8
House-furnishings	1.8	1.8	3.2	0.4	0.5	3.3
Miscellaneous	6.2	6.1	7.3	5.7	6.9	6.2
All commodities	100.0	100.0	100.0	100.0	100.0	100.0

*Certain of the differences between years in the percentage distributions of weights among the groups in the Bureau of Labor Statistics index are due to changes in the constituent items. This is particularly true of building materials and house-furnishings.

TABLE 86

PERCENTAGE DISTRIBUTIONS OF WEIGHTS AMONG COMMODITY GROUPS
FEDERAL RESERVE BOARD CLASSIFICATION

(1) Commodity group	(2) Percentage of total weight Present study	(3)	(4)	(5) Percentage of total weight Federal Reserve Board classification of U. S. B. of L. S. data	(6)
	1890–1902	1902–1913	1913–1926	1913	June 1923
Crops	13.8	13.5	12.6	12	12
Animal products	14.2	13.9	11.5	15	12
Forest products	3.8	3.7	3.4	3	4
Mineral products	14.7	14.3	13.0	11	16
Total raw materials	46.5	45.4	40.5	41	44
Producers' goods	12.9	14.5	19.4	21	18
Consumers' goods	40.6	40.1	40.1	38	38
All commodities	100.0	100.0	100.0	100	100

[1]The classification of the Bureau of Labor Statistics is that employed by that Bureau prior to the recent revision of its index number of wholesale prices.

The relative importance of the different classes when no weights are used is indicated in the following tables, showing the distribution, among the various groups, of the commodities employed in the present study.

TABLE 87

Distributions of Commodities, by Groups

United States Bureau of Labor Statistics Classification

(1) Commodity group	(2) 1890-1902 No. of commodities	(3) 1890-1902 Per cent	(4) 1902-1913 No. of commodities	(5) 1902-1913 Per cent	(6) 1913-1926 No. of commodities	(7) 1913-1926 Per cent
Farm products	24	12.3	24	11.7	55	14.1
Foods	36	18.5	37	18.1	92	23.5
Cloths and clothing	38	19.5	44	21.5	65	16.6
Fuel and lighting	12	6.1	12	5.9	15	3.8
Metals and metal products	19	9.7	21	10.2	37	9.5
Building materials	21	10.8	22	10.7	32	8.2
Chemicals and drugs	11	5.6	11	5.4	39	10.0
House-furnishings	21	10.8	21	10.2	31	7.9
Miscellaneous	13	6.7	13	6.3	25	6.4
All commodities	195	100.0	205	100.0	391	100.0

TABLE 88

Distributions of Commodities, by Groups

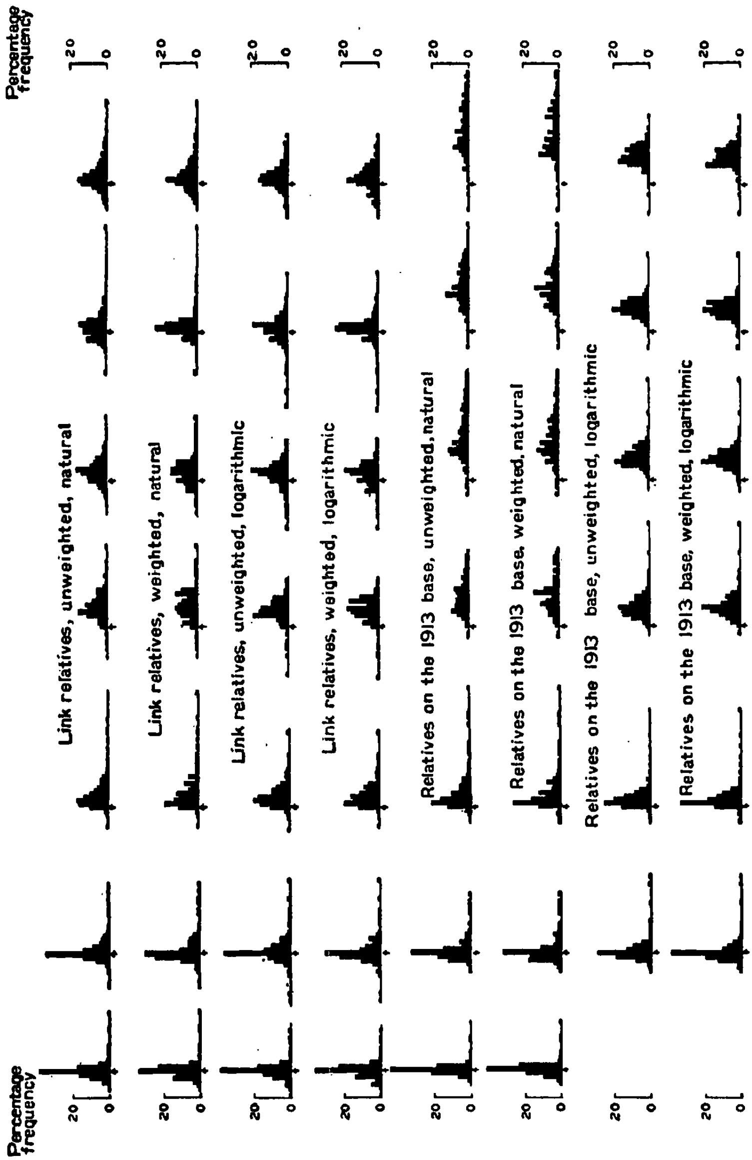
Federal Reserve Board Classification

(1) Commodity group	(2) 1890-1902 No. of commodities	(3) 1890-1902 Per cent	(4) 1902-1913 No. of commodities	(5) 1902-1913 Per cent	(6) 1913-1926 No. of commodities	(7) 1913-1926 Per cent
Crops	9	4.6	9	4.4	21	5.4
Animal products	11	5.7	11	5.4	21	5.4
Forest products	7	3.6	7	3.4	11	2.8
Mineral products	17	8.7	17	8.3	30	7.6
Total raw materials	44	22.6	44	21.5	83	21.2
Producers' goods	51	26.1	54	26.3	112	28.7
Consumers' goods	100	51.3	107	52.2	196	50.1
All commodities	195	100.0	205	100.0	391	100.0

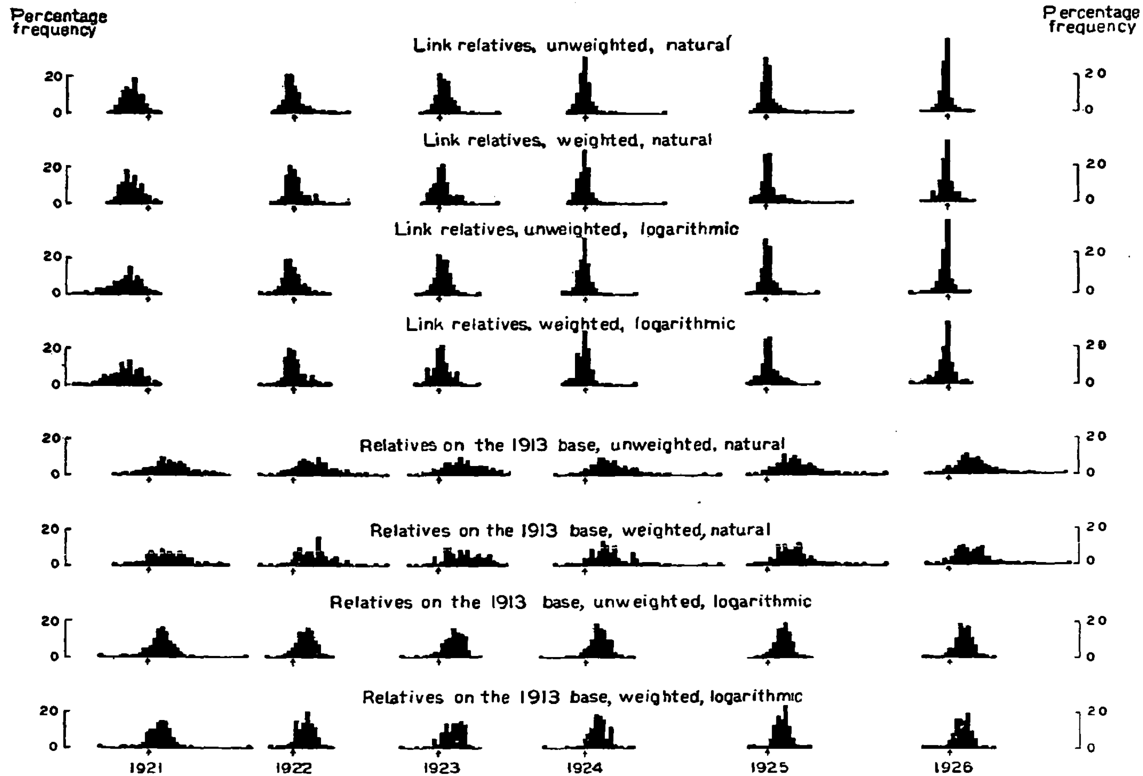
3. Significance of Frequency Distributions of Price Relatives

A bird's-eye view of price changes between 1913 and 1926 is afforded by the graphs of distributions of price relatives which are plotted in Figure 21. These graphs afford, also, a means of com-

FIGURE 21
COLUMN DIAGRAMS SHOWING FREQUENCY DISTRIBUTIONS OF PRICE RELATIVES[1]
(1914-1920)



COLUMN DIAGRAMS SHOWING FREQUENCY DISTRIBUTIONS OF PRICE RELATIVES[1]
(1921-1926)



[1] The number of observations included in these distributions varies from 387 to 391.
Class intervals are the same for unweighted and weighted distributions. They are not the same for distributions of link and fixed base relatives. The slight differences between the link and fixed base distributions for the year 1914 are due to this fact.
The location of the 100 point on the x-scale (the 2.00 point when logarithms are employed) is indicated in each diagram by an arrow.

paring visually the distributions secured from fixed base and link relatives, from weighted and unweighted relatives, from relatives in natural and relatives in logarithmic form. Class frequencies have been expressed in percentage form throughout, so that direct comparison of all distributions is possible.

It is clear that these distributions possess features common to all collections of quantitative data relating to social and biological phenomena. From the host of individual price changes, which seem erratic and unrelated when viewed as separate entities, have been secured distributions with distinct and definite characteristics. The fact that something approaching a common type of distribution seems to prevail among the groups represented in these charts justifies the application of a common method in measuring their distinctive features. But more pronounced than any general resemblances are the differences between the distributions for different years and for different forms of the original data. It is these differences which are matters of interest.

The measures describing the distributions which are graphically portrayed in this chart, and those relating to distributions for earlier years, enable the nature and magnitude of these differences to be precisely determined. Having these measures it is possible to consider several questions which are of considerable importance in connection with the theory and practice of index number construction, and which have a direct bearing upon the problem of price stability. The following specific questions may be raised:

1. What economic significance attaches to changes in the four basic measures descriptive of frequency distributions of price relatives (i. e. the measures of central tendency, of dispersion, of skewness and kurtosis)? What light is cast by these measures upon the problems outlined above, relating to changes in the general price level and to the internal stability of the price system?
2. Are there significant differences between the different types of frequency distributions? Do distributions based upon unweighted price relatives differ materially from those based upon weighted relatives? Do logarithmic distributions differ materially from natural distributions? Do distributions composed of link relatives differ materially from those constructed from fixed base relatives? Which type of distribution appears to be most stable? Which type of distribution

gives the most reliable measures of central tendency and dispersion?
3. Is there any tendency for the distributions of price relatives to approach a standard type of distribution (e. g. the normal or Gaussian type)?
4. Is there any evidence of a change in type over the period of time covered in the present study? Are there significant changes in type with changing business conditions?
5. Does a study of these distributions afford evidence as to the stability or instability of distributions of price relatives? Does it afford evidence as to the homogeneity or heterogeneity of distributions of price relatives?

In attempting to answer these questions methods of describing frequency distributions in detail must be used. The measures utilized in the present study are those developed by Karl Pearson in connection with his comprehensive system of frequency curves. Such measures afford materials for comparing and classifying frequency distributions, and reveal the distinctive characteristics of individual distributions. Detailed accounts of the method followed in computing these measures are available in Pearson's original memoirs[1] and in various other sources. The measures secured from the application of these methods are given in Appendix Tables XIX-XXVII. They are considered in detail in the following sections of this chapter.

§Note on the Description of Frequency Distributions

The problems which called forth Pearson's measures arose from the failure of the normal curve of error to describe the actual distributions found in experience. Marked and persistent departures from this standard type were found in every scientific field. If one were not to assume that all such departures resulted from lack of homogeneity in the data, an assumption that seemed quite unjustified, it was necessary to recognize the existence of non-Gaussian types of distributions, and to develop methods of distinguishing these types from the normal form and of describing and graduating them. This is a problem which has been attacked in various ways. Pearson's solution is the one which has been worked out in greatest detail and is most widely employed.

[1] "Skew Variation in Homogeneous Material," *Phil. Trans. of the Royal Society of London*, Vol. 186, A (1895), pp. 343-414; Supplement, Vol. 197, A (1901), pp. 443-459; Second Supplement, Vol. 216, A (1916), pp. 429-457.

"Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson," A Rejoinder. *Biometrika*, Vol. IV (1905), pp. 169-212.

In a particular case the general problem reduces itself to two distinct questions:

1. Is the distribution in question significantly different from the normal type? In any concrete instance some difference from the normal is to be expected, as a result of chance fluctuations of sampling, hence a test of the significance of such differences is necessary.
2. If the distribution is significantly different from the Gaussian type, is it possible to place it in some standard classification and to employ a generalized method of graduation similar to that employed when a normal curve is fitted?

In respect to deviations from the normal type, Pearson set forth the following as the chief physical differences between actual frequency distributions and the Gaussian theoretical distribution:

1. The separation between the mode, or position of maximum frequency, and the average or mean character.
2. The ratio of this separation between mean and mode to the variability of the character, a quantity termed the *skewness.*
3. A degree of peakedness (or, viewed inversely, of flat-toppedness) which is greater or less than that of the normal curve. This attribute relates to the *kurtosis* of the curve.

Pearson developed measures for these characteristics (the modal divergence, the skewness and the kurtosis) each of which would be zero for the normal type. He determined, also, the probable errors of these measures, by means of which it is possible to determine whether a given value of any measure represents a significant departure from the Gaussian type. The first of the preceding questions may, therefore, be readily answered for a given distribution.

So far the results are purely negative, if the normal curve fails to fit the given distribution. The next step is the development of a technique for describing and graduating the non-Gaussian distributions which are met with so frequently in statistical practice. This Pearson has done. The process of evaluating the differences between an actual frequency distribution and the Gaussian theoretical type involves a knowledge of the dispersion, the skewness and the kurtosis of the given distribution. These same three measures in slightly different form furnish criteria which serve to define the curve type, whether it be Gaussian or non-Gaussian. The other specific attributes of the curve desired in a given case (its position on the x-scale and its size) are fixed by the value of the mode, or mean, and the total frequency.

The details of the methods employed in graduating these distributions do not concern us in the present study, but considerable interest attaches to the criteria of curve type and the information they yield.

Five such criteria have been computed for each of the price distributions listed above. These are

$$\beta_1 = \frac{\mu^2_3}{\mu^3_2}$$

$$\beta_2 = \frac{\mu_4}{\mu^2_2}$$

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

$$\kappa_2 = \frac{\beta_1\,(\beta_2 + 3)^2}{4\,(4\beta_2 - 3\beta_1)\,(2\beta_2 - 3\,\beta_1 - 6)}$$

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

where the letters μ_1, μ_2, μ_3 and μ_4 designate the first four moments about the mean, adjusted, where necessary, by the application of Sheppard's corrections.

Differences between Gaussian and non-Gaussian curves are defined by the following measures:

$$\eta \text{ (Kurtosis)} = \beta_2 - 3$$

$$\chi \text{ (Skewness)} = \frac{\sqrt{\beta_1}\,(\beta_2 + 3)}{2\,(5\,\beta_2 - 6\,\beta_1 - 9)}$$

$$d \text{ (Modal divergence)} = \chi . \sigma$$

III Changes in the Level of Wholesale Prices

The immediate purpose of the present investigation is not the measurement of changes in the level of prices, a subject which has been discussed extensively elsewhere. But in a study of the behavior of prices in combination some attention must be given to such general price changes, for these constitute one important aspect of group behavior.

1. Comparison of Index Numbers

The column diagrams which are shown in Figure 21 differ from year to year in many ways—in the location of the point of maximum concentration, in the degree of dispersion, in the direction and degree of skewness, and in peakedness. Our present concern is with the shifts in the central tendency from year to year, as measured by the changing values of the annual averages. Prices as a whole drift upward or downward, and the changing position on the x-scale of the point of maximum concentration is an indication of the direction and degree of this drift. It is possible to follow this drift on the charts by noting the varying distances between the central ten-