Volume Title: The Behavior of Industiral Prices

Volume Author/Editor: George J. Stigler and James K. Kindahl

Volume Publisher: UMI

Volume ISBN: 0-87014-216-X

Volume URL: http://www.nber.org/books/stig70-1

Publication Date: 1970

Chapter Title: APPENDIX B The Construction of the NBER Index Numbers

Chapter Author: George J. Stigler, James K. Kindahl

Chapter URL: http://www.nber.org/chapters/c3320

Chapter pages in book: (p. 101 - 104)

APPENDIX B

# The Construction of the NBER Index Numbers

This section discusses the transformation of the edited data on individual commodities from the various informants into the NBER indexes for these commodities and commodity classes.

All indexes were constructed in such a way as to make the absolute level of each price series irrelevant. This decision was dictated by the nature of the data, and all subsequent interpretations and computations had to be made compatible with this. (The reason for this requirement is explained more fully below.) Different price series could be combined either by assuming that price series tended to differ by a constant number of dollars per physical unit of product, or by assuming that price series tended to differ by a constant percentage amount. It is not hard to find appropriate examples for each assumption. Price differences due only to freight differences, or due to quantity discounts expressed in dollars independent of the price per unit, would lead to absolute differentials. Price differences stemming from discounts of the more common variety, expressed as a percentage of the same base price, would represent percentage differentials.

Ultimately the choice of percentage differentials was forced upon us by a combination of factors. Some data were available only in the form of index numbers; certain firms insisted upon expressing their data this way, so that absolute levels of price were never disclosed. In other cases, we found the degree of product heterogeneity within some categories sufficiently great that we were reluctant to compare the dollar values of various items. These and other considerations argued for the

calculation of indexes which preserve percentage differences rather than absolute differences between observations within a given series. The choice was reinforced by both computational ease and our intuition of the prevalence of parallel price movements of similar commodities in percentage terms.

Since we chose to preserve percentage or relative differences in individual series over time, it follows that the average taken of individual series at each point in time must be a geometric mean, not an arithmetic one. The only information contained in an individual firm's data, once it has been subject to multiplication by an arbitrary constant, is the relative change from month to month, so a meaningful index must preserve this change; and the geometric mean does. All computations were therefore performed as arithmetic operations on logarithms of the original data.

Since neither the time span covered nor the frequency of observation was the same for all series, it was necessary to adjust individual price series. The procedure was in two main steps. For each commodity group, the first step was to select out all those series which were either monthly data to begin with, or became so as a result of editing—mostly as a result of recording contract prices in each month of the period covered. An index of such monthly data was then constructed. Second, each series which was not monthly was interpolated, for missing observations, using the previously computed monthly index as a related series, where this was possible, and linear interpolation otherwise. The set of irregular or broken series thus interpolated into monthly series was then combined into an index of nonmonthly prices; and finally, the two monthly and nonmonthly indexes were combined into a single index weighted by the number of series in each. The mechanics of these operations are discussed in some detail below.

## CONSOLIDATION OF MONTHLY SERIES INTO A SINGLE INDEX

Since the various monthly price series begin and end in different months, they are combined by averaging the monthly relatives (first differences in logarithms) for each adjacent pair of months and linking these relatives to form a continuous series. The price index is simply the antilogarithms of the resulting series, with 1964 equal to 100.

## INTERPOLATION OF NONMONTHLY SERIES

All series were interpolated so that they had no gaps in the months between the first and last observations. Interpolation in this case, for purposes of constructing an average, is an eminently reasonable procedure. We cannot afford to discard all the information contained in nonmonthly data. Milton Friedman has shown that an efficient interpolation procedure commends the use of an interpolating series that is well correlated with series to be interpolated; and, furthermore, that the interpolating series be followed more closely, the better it is correlated with the series to be completed.[1]

We adapted the Friedman procedure with the price data in logarithmic form. After the index of monthly series was computed, it was used as a related series to interpolate each of the nonmonthly series in turn. From all dates at which the monthly index and the nonmonthly series both had observations, the regression coefficient was computed. The initial step was to interpolate linearly in the nonmonthly series (thus imparting constant rates of change between observations), then to transfer the appropriate multiple (the correlation coefficient multiplied by the ratio of standard deviations is the multiplier) of the corresponding deviation of the monthly index from its constant rate of change trend to the series to be interpolated. This second component was superimposed on the linear interpolation already performed. Each nonmonthly series was so interpolated using the same "related" series; however, each such series had a regression coefficient of its own and, hence, the proportion of the variation in the monthly index which was imposed on the nonmonthly series differed. In the interpolation of one particular series, no attempt was made to use information from other nonmonthly series. The computational problems involved would have become formidable. When irregular price series were available for periods in which no monthly prices were available, we used linear interpolation.

After interpolation each of the nonmonthly series became a monthly series. In combining these series into a continuous index of nonmonthly prices, the differences in beginning and ending dates of individual series were handled exactly as with the true monthly series.

[1] *The Interpolation of Time Series by Related Series,* Technical Paper No. 16, New York, NBER, 1962.

For each commodity, the nonmonthly and monthly indexes were combined into a single index, which thereby utilized all available information, in the usual manner, i.e., using as weights the number of series. This formed the fundamental NBER index for each commodity.

## COMBINATION OF INDEXES

The industrial price indexes in Chapter 6 are calculated with a conventional Laspeyres formula, using the BLS value of shipment weights for 1958. The BLS indexes are also recalculated, employing only commodity price indexes for which the NB has coverage.

The weighted commodity indexes in Chapter 4 are calculated by a similar procedure, using as weights the average quantities purchased by the individual price reporters in 1960–62.