New research on interarea consumer price differences

An experimental interarea price index project is under way at BLS and has generated indices for several components of the consumer's budget; the aim of the project is to adapt CPI data to facilitate interarea cost-of-living comparisons

Mary F. Kokoski

Interarea comparisons of the cost of living are of interest to researchers and policymakers, as well as to businesses and households making decisions about where they should locate or relocate. It is widely perceived that the cost of goods and services differs across geographic areas. Published Consumer Price Index (CPI) series for various geographic areas measure price changes over time for these areas. An official, consistently designed index of interarea price differences does not exist. Research is under way, however, to investigate the possibility of using the enormous data resources compiled for the CPI to construct a comprehensive index of interarea prices. Such an index, if officially produced, would provide an interarea complement to the CPI.

Among the goods and services included in this experimental index would be those composing the food-at-home component of the consumer's budget. Food at home, which constituted 10.1 percent of the average household budget in December 1989, is also an expenditure of particular interest to policymakers and researchers in the health field. Accordingly, in this article, an experimental interarea price index for food at home is analyzed in the context of past research at the Bureau of Labor Statistics on geographic differences in prices.

Mary F. Kokoski is an economist in the Division of Price and Index Number Research, Bureau of Labor Statistics.

Interarea price measurement

The first official measure of interarea differences in the cost of living was the standard budget of the Family Budgets program of the BLS, developed in the 1940's. Under this program, a predetermined standard of living was translated into a set of hypothetical baskets of goods and services, and the costs of these baskets were compared across geographic areas. The standard budget embodied an "income requirements approach," by assuming that families with a given demographic profile had the same utility function across goods and services, leisure, and working conditions.[1] It did not, however, take an identical basket of goods and services as a common denominator to compare in all areas. Variations in the specific items included, as well as in the quantities purchased, were therefore implicit in the interarea comparison. As Mark Sherwood pointed out in 1975, an interarea price index produced from these budget data would serve as a measure of differences in the cost of living only under the assumption that a given family would be indifferent with respect to all the various baskets.[2] By comparing index values from the Family Budgets program with those generated by a fixed market basket for all areas, he found that variations in specification of items and in quantities did have a large effect on the interarea comparisons for food and transportation.[3]

Although the Family Budgets program was phased out in the 1970's, price information continues to be collected in many metropolitan and nonmetropolitan areas in support of the Consumer Price Index and the Average Prices program of the BLS. Unfortunately, these price data are not readily amenable to interarea comparisons. First, the Average Prices program includes only selected food items in its sample, and there are variations in the brand and quality of items

across areas.[4] Second, average retail prices for food are reported officially only at the U.S. and regional levels. Third, the price sample used for the CPI, although extremely large, is intended to facilitate intertemporal comparisons and is based on a probability-sampling methodology. This means that the specific items priced in each outlet in each area are those which consumers in that outlet and area are likely to purchase. Thus, while a 10-ounce jar of Folger's crystals will represent the instant coffee category in a grocery in Milwaukee, a 2-ounce jar of house-brand instant coffee will represent the same category in a grocery in Denver. This diversity of market baskets makes direct comparison of prices across areas problematic.

An experimental interarea price index

Recent research has demonstrated that interarea comparisons can be made using the CPI price quotes as a resource.[5] To accomplish these comparisons, a methodology is adopted which accounts for the diversity of items in the sample in a statistically explicit fashion. Based on the Country-Product-Dummy method, used in making international output comparisons,[6] this approach employs a hedonic regression, with each specific item's price as the dependent variable and three sets of independent dummy variables: one describing the item's physical characteristics (such as weight and color), a second defining the type of outlet in which the item is sold (for example, a chain supermarket), and a third defining the geographic area in which the item is sold (for instance, Milwaukee). When the dependent variable is expressed in logarithmic form, the coefficients of the area dummies can be interpreted as bilateral interarea price indices, with one area arbitrarily chosen as the reference area by excluding its dummy variable from the regression sample.

When aggregated across commodities, these bilateral indices, while useful for many applications, are not, in general, transitive. For example, the product of the bilateral index for Baltimore to New York with that of New York to Philadelphia will not, in general, equal the direct bilateral index of Baltimore to Philadelphia. Also, alternative choices of reference area will result in different index values for all areas. For instance, the price differential for Baltimore versus Milwaukee may be different when Philadelphia is taken as the reference area than when New York is chosen as the reference area. To overcome these difficulties, the bilateral indices can be adjusted by a method of constructing a transitive, multilateral set of index values described initially by O. Elteto, P. Koves, and B. Szulc.[7] This method, known as the EKS approach, takes the geometric mean of all areas in the sample as the reference "area."

The construction of an EKS index requires four basic steps. First, hedonic regressions are run at the most disaggregate level possible of item classification, the entry-level item. Second, the regression coefficients are weighted up, to construct interarea price relatives at a more aggregate level, the item stratum level. The weights used are the preliminary cost weights from the Consumer Expenditure Survey. Third, the price relatives are used to construct a set of bilateral Törnqvist indices that are aggregated over item strata to the level of aggregation desired (for example, all food at home). Fourth, the geometric mean of the bilateral Törnqvist indices is obtained. The value of this mean for an area i is the EKS index value for that area.

Formally, the Törnqvist (transcendental logarithmic) price index is defined as

$$(1)\quad \ln P_{ij}^{T} = \frac{1}{2}\sum_{n}\left(s_i^n + s_j^n\right)\cdot \ln\left(\frac{P_i^n}{P_j^n}\right)$$

$$= \frac{1}{2}\sum_{n}\left(s_i^n + s_j^n\right)\cdot\left(\bar{b}_i^n - \bar{b}_j^n\right) \qquad i, j = 1, \ldots, M$$

where the summation is over the item strata composing the aggregate of interest, s_k^n is item stratum n's share of expenditures for that aggregate in area k, $\ln\left(P_i^n / P_j^n\right)$ is the difference between the item stratum-level relative interarea prices for area i and area j for item stratum n, and $\bar{b}_i^n$ is the average of the hedonic regression coefficients for item stratum n in area i. The EKS index δ_i is then defined as a weighted geometric mean of the Törnqvist bilateral price indices for area i relative to every other area j:

$$(2)\quad \ln \delta_i = \sum_{j=1}^{M}\left(\frac{p_j \cdot q_j}{\sum_l p_l \cdot q_l}\right)\ln P_{ij}^{T}$$

Here, $p_j q_j$ is expenditures in area j, so the weights are area j's share of expenditures in all areas. The multilateral indices δ_i, then, represent the price level in area i relative to all other areas.

The EKS method has been applied to the food-at-home component of the CPI to derive a multilateral interarea index for the 44 geographic areas that have published CPI series. These consist of 32 self-representing areas[8] and 12 areas that are aggregates of smaller Standard Metropolitan Statistical Areas clustered by census region and population size class. The survey period July 1988 through June 1989 was chosen to provide the sample of price quotes. After a screening of the data for missing information or errata that could not be corrected, the final sample comprised 53,459 observations. The regressions

were run at the level of the entry-level item, and the resulting coefficients were weighted up to obtain a set of price relatives for the next level of aggregation, the item stratum level, using the expenditure weights from the Consumer Expenditure Survey.[9] The EKS index thus derived covers the aggregate of all food at home, as defined by the CPI.[10]

Index values for selected cities are presented in the first column of table 1 as the variable IAP (for "Interarea Price"). Unfortunately, computation of variance estimates for these indices is not straightforward, a result of imputation procedures required in the aggregation from entry-level item coefficients to the item stratum level. Research is under way to enhance the procedure.

Interarea indices for food at home

Although they cover different time periods and were calculated by different methodologies, it is interesting to compare the interarea index values from the hedonic EKS method with the values obtained by Sherwood with a fixed market basket. Similarities may indicate consistent patterns of price differences across cities and also lend some support to the usefulness of the hedonic method in statistically controlling for differences in item characteristics. Thus, in table 1, column 2 shows the Sherwood index values for the year 1973, the reference year of his published study, and column 3 gives a set of "updated Sherwood" index values for the year 1988, more comparable to the time period of the IAP values. Both of these Sherwood series are based on his fixed-weight index.[11]

To obtain Sherwood index values comparable to the IAP results, it was first necessary to identify geographic areas that were the same in both studies. These are the 25 cities in table 1. Taking these cities as the relevant sample, the index values for both Sherwood and IAP were "reweighted" so that the index base was the mean index value for this sample for each type of index. That is, the reference for each column of indices is the mean of that column of index values across all cities in table 1. Because of the 15-year time difference between the two studies, a set of "updated Sherwood" indices was constructed by moving the reweighted 1973 Sherwood index values forward by each city's CPI for 1973 to 1988. That is, each reweighted Sherwood index value in column 2 was multiplied by a price relative of the 1988 food-at-home CPI to the 1973 food-at-home CPI, where the CPI values were specific to each respective city. The resulting index values were then reweighted to give them a base value of 100 and are presented in column 3.

A comparison of the IAP values to the 1973 Sherwood values indicates a different pattern of cities above and below the mean of interarea prices for food at home. However, when the Sherwood index is updated by the CPI, the pattern is quite similar to that of the IAP index. Only two cities, Boston and Kansas City, deviate from the common pattern. In the absence of variances and, thus, tests of statistical significance, this observation would appear to provide some support for the hedonic method when identical market baskets cannot be compared across areas. The difference between the 1973 Sherwood values and the "updated Sherwood" values suggests that interarea price differences are not consistent over time.

The "updated Sherwood" 1988 and IAP 1988–89 pattern of price differences shows generally higher prices for cities in the south and more cities with below-average prices in the north central region of the United States. In all index series, the high price outliers are Anchorage and Honolulu, which would be expected on the basis of heuristic evidence. The range of IAP index values is from 93.4 to 106.8 for cities in the continental United States; for the "updated

Table 1. **Index values for food at home, selected cities**

[Average=100]

City	IAP,[1] 1988–89	Sherwood, 1973	Updated Sherwood, 1988
Northeast			
Boston	99.3	102.2	102.6
Buffalo	96.6	96.2	98.8
New York/Northeastern NJ	105.1	103.2	112.0
Philadelphia	102.1	99.2	102.2
Pittsburgh	93.3	95.2	91.0
North Central			
Chicago	102.2	102.2	102.3
Cincinnati	106.2	98.2	106.2
Cleveland	96.4	95.2	96.0
Detroit	98.3	102.2	98.2
Kansas City	101.3	100.2	99.2
Milwaukee	95.9	93.3	94.8
Minneapolis/St. Paul	93.4	96.2	98.7
St. Louis	105.7	100.2	103.7
South			
Atlanta	106.8	103.2	104.1
Baltimore	105.1	101.2	101.6
Dallas	103.2	94.2	100.5
Houston	102.4	99.2	110.2
Washington	104.1	103.2	112.8
West			
Denver	96.5	95.2	90.9
Los Angeles	100.6	93.3	100.0
San Diego	97.3	91.3	98.6
San Francisco	102.4	98.2	107.4
Seattle	104.3	100.2	100.7
Noncontinental			
Honolulu	139.0	116.1	141.3
Anchorage	125.7	121.0	124.7

[1] IAP = interarea price index.

Sherwood" index, the range is somewhat broader, from 91.0 to 112.9, although the statistical significance of these differences has not been assessed.

Future research

At present, the experimental interarea price index project has provided indices for food at home,[12] professional medical services,[13] and housing.[14] Indices for private transportation commodities, apparel, and household fuels and utilities are currently in the process of being formed. When these are completed, coverage will have been extended to about 85 percent of the average household's budget for consumption.[15] The results thus far are consistent with *a prior* expectations of intercity price differences and with information from other sources.

The component indices will be aggregated into an all-items index for those goods and services in the project sample. This index, and the information obtained by constructing each of the component indices, could provide useful guidelines for adapting the CPI sample of price information to facilitate interarea as well as intertemporal comparisons. □

Footnotes

ACKNOWLEDGMENT: Thanks are due to Kimberly Zieschang for his helpful comments and to Joseph Pavalone for assistance with research.

[1] Mark K. Sherwood, "Family budgets and geographic differences in price levels," *Monthly Labor Review*, April 1975, pp. 8–15.

[2] Sherwood, "Family budgets," p. 9.

[3] *Ibid.*, p. 8.

[4] Floyd A. Rabil, "Average retail food prices: a brief history of methods," *Monthly Labor Review*, November 1984, pp. 52–53.

[5] See L. Blanciforti, *Price Measurement for Interarea Comparisons*, unpublished manuscript (Bureau of Labor Statistics, November 1986); D. F. Primont and Mary F. Kokoski, *Comparing Price across Cities: A Hedonic Approach*, BLS Working Paper 204 (Bureau of Labor Statistics, May 1990); and D. F. Primont and Mary F. Kokoski, *Differences in Food Prices across U.S. Cities: Evidence from CPI Data*, BLS Working Paper 209 (Bureau of Labor Statistics, November 1990).

[6] See I. B. Kravis, "Comparative Studies of National Incomes and Prices," *Journal of Economic Literature* 22 (1984): 1–39; I. B. Kravis, Z. Kennessey, A. Heston, and R. Summers, *United Nations International Comparison Project: Phase One, a System of International Comparisons of Gross Product and Purchasing Power* (Baltimore, the Johns Hopkins University Press, 1975); and R. Summers, "International Comparisons with Incomplete Data," *Review of Income and Wealth*, March 1973, pp. 1–16.

[7] See W. E. Diewert, "Index Numbers," in *The New Palgrave Dictionary of Economics*, vol. 2, edited by J. Eatwell, M. Milgate, and P. Newman (New York, Stockton Press, 1987), pp. 767–80; and L. Drechsler, "Weighting of Index Numbers in Multilateral International Comparisons," *Review of Income and Wealth*, March 1973, pp. 17–34. The method was generalized in D. W. Caves, L. R. Christensen, and W. E. Diewert, "Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers," *Economic Journal*, March 1982, pp. 73–86.

[8] A self-representing area is an area represented in the CPI by a sample of price quotes from that area alone. For further details on the relationship between self-representing areas and Standard Metropolitan Statistical Areas, see *BLS Handbook of Methods*, Bulletin 2285 (Bureau of Labor Statistics, 1988), pp. 161–162.

[9] These weights are the same preliminary cost weights used for the CPI; see *BLS Handbook of Methods*, p. 168.

[10] For more details, see Primont and Kokoski, *Differences in Food Prices*.

[11] Index 4 in Sherwood, "Family budgets," p. 10.

[12] Primont and Kokoski, *Differences in Food Prices*.

[13] Primont and Kokoski, *Comparing Price*.

[14] B. Moulton, *Interarea Indexes of the Cost of Shelter Using Hedonic Quality Adjustment Techniques*, unpublished manuscript (Bureau of Labor Statistics, December 1989).

[15] See *CPI Detailed Report for December 1988* (Bureau of Labor Statistics, 1989).