# Measuring Retail Food Price Variation:
# Does the Data Source Matter?

**Ephraim Leibtag**
**Food Markets Branch**
**Economic Research Service**
**USDA**
eleibtag@ers.usda.gov

# Measuring Retail Food Price Variation: Does the Data Source Matter?

Can Americans afford a healthful diet? One major determinant of affordability is price. In this context food prices play an important role in consumers' food choices. Since food prices can vary across markets and across demographics and locations within markets, this research focuses on estimating the extent to which price variation exists on a number of levels. Since food choices are made at the local level, a model of consumer food choice should use disaggregated prices measured to best match prices available to a given consumer at the time of the actual purchase decision.

The main area of focus in this paper is variation across regions and over time in the U.S. as estimated by different food price data sets. Once differences in food prices have been estimated, the variation in prices can be used in a food choice model to test whether or not prices matter in the context of what foods people choose to purchase and consume. Presumably, these food consumption choices will affect the health and nutrition outcomes of the U.S. population.

There are a variety of potential sources for food price data, but it is important to compare the relative strengths and weaknesses of the data sources in order to estimate the magnitude of the variation in average prices from the different sources. No one data source can account for all potential measurement problems, but by estimating average food prices for similar, or even

identical products, one can test for whether there are significant differences in average food prices based on the data source.

This paper uses four distinct sources for food price information for 2005. Two of the price sources are from Nielsen data: Scantrack and Homescan, while the third source is a survey of prices collected by the Council for Community and Economic Research (formerly known as ACCRA) to compare costs of living across U.S. markets, and the final data source is the Average Price data collected by the Bureau of Labor Statistics (BLS) as part of the Consumer Price Index data collection process.

The Nielsen Scantrack data contain weekly sales volume and quantities sold for any product labeled with a Universal Product Code (UPC) and include many food items in grocery and other retail food stores for a sample of over 12,000 retailers. These retailers include most of the major chains in the U.S. and therefore account for the majority of food-for-home consumption sales in the country. Weekly average prices are calculated for selected UPCs, as well as broader food categories, by dividing the dollar sales volume by the number of units sold. The deviation from the actual transaction price is small and depends on the extent to which a given retailer offers the same product at more than one price in a given week (for example, one price for those using a shopper loyalty card and one price for those that do not) and the extent to which consumers use

coupons or other consumer-specific discounts at the time of purchase.  The strength of this data source is the broad coverage of many products within a given category as well as coverage of most of the major food retailers in many of the major U.S. markets.  However, these data are far from perfect, as some retailers are not included in the Nielsen Scantrack sample and products that do not have UPC codes are not tracked in the data.  This is especially troublesome for analyses of fresh produce and meats that often do not have UPC codes given that many are not produced and branded by the major manufacturers that use the UPC system.

The second data source from Nielsen is the Homescan Consumer Panel. This data set tracks actual prices paid by consumers on a weekly basis from all of the stores that a given shopper chooses to purchase food from.  These data also cover non-UPC coded products, so that both weaknesses of the Scantrack data are improved on.  However, the Homescan data are a much smaller sample of food prices as compared to the Scantrack data and allow only for the observation of the transaction price for a particular household. The Homescan data are collected with the goal of representing U.S. census demographics, but this methodology assumes that standard consumer demographics are a main determinant of food shopping behavior when, in fact, other consumer characteristics not captured or accounted for in the data could be larger determinants of shopping behavior.

The third data source, the ACCRA data, contains the least amount of price information in terms of food coverage, but has the broadest geographic coverage of all of these data sets, allowing for a large number of cross-market comparisons. In addition the ACCRA data contains some information on food-away-from-home prices that is not included in either of the Nielsen data sets. Finally, the BLS Average Price (AP) data is valuable as a comparison set of prices since it is collected using the best documented, and most statistically sound, price collection methodology of the four data sets. Unfortunately, the AP data only covers a limited number of food product categories, as compared to the Nielsen data sets and the data are only available at the regional and national level.

**A Closer Look at Egg Prices**

Given initial data limitations, it is important to first focus on food products that are available to price in all four data sets. Retail egg prices are a good product to compare across data sets in order to highlight the extent to which prices vary since eggs are a relatively homogeneous product both across stores and regions of the country. Even for this relatively easy-to-compare product, though, not all of the data sets can be directly compared since product coverage varies. Two sets of comparisons are made: 1) the average price of a one dozen eggs (all types) as estimated using the Nielsen Homescan and Nielsen Scantrack data sets and 2) the average price

of large eggs as estimated using the Nielsen Homescan, BLS AP, and ACCRA data sets. Disaggregated Homescan allows for calculations of both the overall average egg price and the specific egg-size types, while the other three data sets contain information in a more aggregate form that limits the comparisons that can be made.

Looking first at the 2005 overall egg prices, it is interesting to note that both Homescan and Scantrack prices show the same trend over time (by quarter) and across regions (table 1). Both data sets show higher prices for eggs in the first and last quarters of the year, with prices at least 10 percent lower in the middle two quarters. Both of these data sets also show egg prices are at least 10 percent lower in the Midwest and South regions as compared to the East and West regions. How then do these data differ? The main difference that persists both across time and regions is the average price level. Homescan prices are lower in 13 of the 16 quarter-by-region price estimates with prices about 4.5 percent lower in the Homescan data. This is consistent with the fact that Homescan include price observations from a wider array of store types and that the main category missing from Scantrack data are some discount stores, such as Wal-Mart and Costco.

The second set of comparisons focuses on just large-sized eggs as this is the egg variety that is tracked by both the BLS AP and ACCRA data sets (table2). The disaggregated nature of the

6

Homescan data allows for a separate calculation of just large-sized eggs to compare with the other two data sources.  Interestingly, more variation in prices is observed in this comparison even though the product in this case is more narrowly defined.  The average price of one dozen large eggs is 15 to 20 percent lower in the Homescan data as compared to the BLS AP and ACCRA data sets.  This is again most likely a function of the coverage differences of these data sets.  The ACCRA data have a limited number of price observations from a self-selected set of stores within a given market, while the BLS AP data reflect a more scientific selection process, but fewer markets included in each regional estimate.  The Homescan data, focusing on actual household purchase behavior shows a lower average price partially as a function of consumer choice to buy food at lower-priced stores.

**Conclusion**

Comparison of regional egg prices show both Midwest and South region prices 10 to 40 percent lower, on average, than East and West region prices, but the variation in prices estimates between both Nielsen data sets, the BLS AP data, and the ACCRA data show how sensitive price estimates can be across data sources.  The trends across time are more consistent across data sets, so that analysis looking just at changes over time may yield similar results regardless of the data set used.  However, if price levels are the focus of research, it is important to recognize that

estimates may be biased from one data source given that it may not be representative of actual prices in a given region.  The preferred approach for future analysis would be to include estimates from multiple data sets to test the robustness of price estimates.  In addition, these results show that national average prices may not be appropriate for the analysis of food choices and that the depth of analysis and matching criteria will guide the researcher to the best price data to use for a given project.

Future research should expand the focus of this paper to additional food products to estimate whether the variation in price estimates is persistent across a variety of food categories or is certain food categories are more uniformly measured across the data sets.  This would inform researchers regarding the best uses of these price data for analysis.

Table 1: Average retail egg price per dozen (all sizes) by region and quarter, 2005

| Region | Data Source | Q12005 | Q22005 | Q32005 | Q42005 |
|---|---|---|---|---|---|
| East | Scantrack (all) | $1.44 | $1.31 | $1.30 | $1.46 |
| East | Homescan (all) | $1.23 | $1.21 | $1.22 | $1.25 |
| South | Scantrack (all) | $1.03 | $0.93 | $0.93 | $1.07 |
| South | Homescan (all) | $1.00 | $0.97 | $0.90 | $1.02 |
| Midwest | Scantrack (all) | $0.96 | $0.86 | $0.88 | $1.02 |
| Midwest | Homescan (all) | $1.01 | $0.83 | $1.07 | $0.94 |
| West | Scantrack (all) | $1.42 | $1.25 | $1.28 | $1.38 |
| West | Homescan (all) | $1.28 | $1.19 | $1.09 | $1.28 |

Table 2: Average retail egg prices (large) by region and quarter, 2005

| Region | Data Source | Q12005 | Q22005 | Q32005 | Q42005 |
|--------|-------------|--------|--------|--------|--------|
| East | ACCRA | $1.64 | $1.64 | $1.73 | $1.70 |
| East | Homescan | $1.16 | $1.11 | $1.16 | $1.22 |
| East | BLS AP | $1.58 | $1.57 | $1.60 | $1.66 |
| South | ACCRA | $1.06 | $0.97 | $0.92 | $0.97 |
| South | Homescan | $1.01 | $0.94 | $0.97 | $1.07 |
| South | BLS AP | $0.98 | $0.84 | $0.88 | $0.98 |
| Midwest | ACCRA | $1.13 | $0.90 | $0.88 | $1.09 |
| Midwest | Homescan | $0.88 | $0.82 | $0.86 | $0.93 |
| Midwest | BLS AP | $1.12 | $1.17 | $1.22 | $1.30 |
| West | ACCRA | $1.76 | $1.65 | $1.76 | $1.76 |
| West | Homescan | $1.30 | $1.25 | $1.35 | $1.43 |
| West | BLS AP | $1.45 | $1.39 | $1.44 | $1.46 |