Getting the Price Right: Analyzing and Comparing Food Prices Over Time and Space

Abigail Okrent
USDA-ERS
Abigail.okrent@usda.gov

Megan Sweitzer
USDA-ERS
Megan.sweitzer@usda.gov

Chen Zhen
University of Georgia
czhen@uga.gov

Anne Byrne
USDA-ERS
Anne.byrne@usda.gov

Mary Muth
RTI, International
muth@rti.org

Selected Paper prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC; July 23-25, 2023

The analysis, findings, and conclusions expressed in this manuscript should not be attributed to Circana (formerly IRI).

Getting the Price Right: Analyzing and Comparing Food Prices Over Time and Space

Abstract

Food price data are instrumental for studying food markets, the consumer food environment, and consumer welfare. The BLS Consumer Price Index (CPI) has been the gold standard for price information but has limited use in food and nutrition policy because (1) the CPI only compares price variation across time and many analyses require price variation across time and geographies; (2) foods in the food at home (FAH) CPI are aggregated into categories that are often too coarse for nutrition analyses. We describe a new data product called the Food-at-Home Monthly Area Prices data product (F-MAP), which provides price information that varies temporally and spatially for foods categorized into 90 categories that closely align with the *Dietary Guidelines for Americans*. We compare the short-run relationships between the F-MAP to CPI FAH indexes at the national level, by Census Region, and for select major FAH subcategories. We find that using retail scanner data in constructing price indices puts downward pressure on measurement of food price inflation compared to the CPI, though effects vary by region and FAH subcategory.

Introduction

Recent food price increases and volatility demonstrate the importance of detailed food price data for studying food markets, the consumer food environment, and consumer welfare. Food-at-home (FAH, i.e., grocery) prices increased 3.5 percent in 2020 and 2021 before increasing 11.4 percent in 2022, the largest annual increase since the 1970's (USDA Economic Research Service, 2023). The Consumer Price Index (CPI) published by the U.S. Bureau of Labor Statistics (BLS) is the principal public source of data on food price changes. The CPI is valuable for its timeliness, coverage of sectors across economy, extensive historical data, and rigorous and transparent methodology. However, the CPI has limitations for use in food policy analysis. Specifically, products tracked by the BLS represent what people purchase, which may not necessarily highlight foods that are encouraged or of interest for nutrition and food security goals. Geographic food price data are also limited, and the CPI is not designed to compare relative prices across areas (BLS, 2023b). The CPI reflects only temporal price variation, which has limited use in research based on cross-sectional or panel data, and spatial price differences have been found to be of particular importance in food policy analysis (Zhen et al., 2018).

Food prices vary across geographic areas (Todd et al., 2010, Gunderson et al., 2022) and affect dietary consumption patterns (Andreyeva, et al., 2011; Afshin et al., 2017), with impacts for diet quality and health outcomes. Spatial variation in food prices also has implications for food assistance programs and program participants. Differences in food prices across areas affect the purchasing power of Supplemental Nutrition Assistance Program (SNAP) benefits (Christensen and Bronchetti, 2020) and the Women, Infants, and Children (WIC) program's fruit and vegetable voucher (Cakir et al., 2018), which are not adjusted for geography or cost of living.

More limited research exists on the effects of food prices on food insecurity in the United States, partially due to limitations on available data. Research using the Quarterly Food-at-Home Price Database found that low-income SNAP households that live in areas with higher food prices were more likely to be food insecure (Gregory and Coleman-Jensen, 2013) and that considerable geographic variation exists in the price of healthy foods (Todd et al., 2011). However, price comparisons using unit values may reflect quality differences in products (Deaton, 1988), and measurements of prices of identical products suggest price differences across areas are largely due to product heterogeneity and variety (Handbury and Weinstein, 2015). Therefore, panel price indices that vary both temporarily and spatially provide a valuable resource for analyzing and comparing trends in food prices and in econometric models of the effects of food policy.

To fill this gap, we created the Food-at-Home Monthly Area Prices data product (F-MAP), which provides food price information comparable over both time and space. The F-MAP includes monthly mean unit values and price indices for 90 food categories across 14 geographic areas. The food categories used in the F-MAP align closely with the *Dietary Guidelines for Americans*, can be mapped to the USDA Thrifty Food Plan categories, and are designed to facilitate food and nutrition research.

The F-MAP and CPI are constructed using different data sources and methods, and in this paper, we compare the F-MAP to the CPI so researchers using the data understand how this new data resource compares to established sources of food price data. We compare price indices from the F-MAP to CPI food-at-home indices at the national level and by Census Region. We examine the variation between and long-term trends in the F-MAP and the CPI and how the underlying source data and index formula contribute to differences between the price measurements for several food categories.

Data

Data for this study comes from the USDA ERS Food-at-Home Monthly Area Prices (F-MAP) product and the BLS Consumer Price Index (CPI) for FAH. This section provides an abridged description of and comparison between the two datasets, presenting sufficient detail for a broad understanding of the data sources and technical understanding of the comparisons made in the present study. Extensive documentation is available for both the F-MAP (Muth et al., 2022) and the CPI (BLS, 2023) for readers interested in more details. [1]

ERS Food-at-Home Monthly Area Prices (F-MAP)

The price measures in the F-MAP include monthly unit values and six price indices for 90 food categories across the four Census Regions and for 10 major metropolitan areas. The prices are constructed from IRI InfoScan retail scanner data, which is a proprietary, commercial data product. InfoScan is a nonprobability sample of approximately 60,000 retailers across the U.S. from various channels including grocery stores, supercenters, club stores, and dollar stores. It includes item-level quantities and revenues, from which prices can be derived, at the weekly level. Roughly half of stores in the InfoScan panel include perishable product sales information for uniform and random weight fresh food products, such as packaged fresh produce and fresh food items sold by the pound or the count (e.g., bulk produce or store-packed meat). The InfoScan data are not a representative sample of retailers, and we apply store-level weights developed by RTI International to weight stores in the InfoScan data to represent the population of food retailers in the U.S. (Muth et al., 2021).

[1] The F-MAP was originally called the Monthly Food-at-Home Price Database or MFAHPD. The name was changed in 2023, but the underlying product remains the same. The F-MAP will be more extensively described in a forthcoming Technical Bulletin from ERS.

IRI product dictionaries provide product characteristics for over 1 million food products, including Nutrition Facts label information and nutrition claims data. Information for perishable products, which may not have detailed nutrition information labels, is also available within the IRI product dictionaries.[2] Using detailed product descriptions (e.g., "whole-grain fruit-flavored breakfast cereal") in the product dictionaries, items are mapped to the 90 ERS Food Purchase Groups (EFPGs), which is a taxonomy of food groups designed for food and nutrition policy-relevant analysis. For example, the EFPGs separate most foods into more and less healthful categories based on the 2015-20 and 2020-25 *Dietary Guidelines for Americans* (U.S. Department of Health and Human Services and Department of Agriculture, 2015), such as whole grains and non-whole grains.

The F-MAP covers several geographies and is initially available for a limited set of years. Data are available in the F-MAP for the four Census Regions and ten large metropolitan areas: Atlanta, Boston, Chicago, Dallas, Detroit, Houston, Los Angeles, Miami, New York, and Philadelphia.[3] Data for the Census Region F-MAP prices include information from retailers that are located in both metropolitan and nonmetropolitan areas. The F-MAP data initially covers the period of 2016─2018, which is the focus of our analysis. However, it will be extended to the 2012─2021 period.

The F-MAP includes several measures of price variation. The simplest price measures are weighted and unweighted unit prices on a per-100-gram basis (mean and standard error). It also

[2] For more information, we refer the reader to Muth et al. (2016) where they detail the IRI datasets and how they can be used for economic research.
[3] The full metropolitan area definitions are as follows: Atlanta-Sandy Springs-Roswell, Georgia; Boston-Cambridge-Newton, Massachusetts; Chicago-Naperville, Elgin, Illinois; Dallas-Fort Worth-Arlington, Texas; Detroit-Warren-Dearborn, Michigan; Houston-The Woodlands-Sugar Land, Texas; Los Angeles-Long Beach-Anaheim, California; Miami-Fort Lauderdale-West Palm Beach, Florida; New York City-Newark- Jersey City, New York and New Jersey; and Philadelphia-Camden-Wilmington, Pennsylvania and Delaware.

includes six weighted price indices (Laspeyres, Paasche, Törnqvist, Fisher Ideal, GEKS—named for Gini [1931], Eltetö and Köves [1964], Szulc [1964]—and CCD—named for Caves, Christensen and Diewert [1982]). The Laspeyres, Paasche, Törnqvist and Fisher Ideal are bilateral price indices, having a fixed base period, whereas the GEKS and CCD are multilateral price indices, having a base period within a non-fixed rolling window as data become available. Essentially the bilateral price indices compare prices within a category only to the base period whereas the multilateral price indices compare prices within a category to the base period across all geographies. All price measures are available for each of the 90 EFPGs across the four Census Regions and for 10 major metropolitan areas. In addition to the price measures, the F-MAP also includes variables for the number of stores, total dollar sales volumes, and total sales quantities in grams by EFPG for each geographic area.

ERS is a principal statistical agency with a mission to provide information for statistical, and not commercial, purposes. While the F-MAP uses proprietary commercial data, the price measures published in the F-MAP are aggregated and do not include specific brands or outlets to protect data confidentiality.

BLS Food-at-Home (FAH) Consumer Price Index (CPI)

The CPI measures the average change in price over time of consumer goods and services. The FAH CPI is calculated using data on prices paid by urban consumers for a representative basket of food items. It is normalized to 100 for a chosen base period, and percent change in the CPI is an estimate of the percent change in the price level over the base period.

The CPI commodities and services survey (CPI survey) collects about 94,000 prices per month for thousands of specific products, including FAH products. FAH prices are collected every month in the CPI survey in person, over the phone, or from web data collected from vendors websites. In-person data collection is the most common and represents roughly two thirds of price data collection, while about eight percent of prices are from store websites. Prices include the retail price and direct taxes for each good, i.e., the cost to the consumer for the item. Selection of stores into the commodity and services survey is based on data from the Consumer Expenditure (CE) Survey, which is also the underlying source data for weighting items priced in the index (BLS 2023). The CPI uses expenditure data from the CE survey to identify items purchased by households and to weight these items in the CPI to reflect their share of consumer spending. Unlike the scanner data used for the F-MAP, both the CE and the commodities and services survey are multistage probability samples, which allow for unbiased estimates of the price indices.

The CPI survey is collected from a geographic sample of U.S. establishments. The CPI survey sample covers 75 urban areas where urban is defined as having at least 10,000 inhabitants. The BLS notes that 93 percent of the U.S. population lives in such areas. Prices are always collected in Chicago, Los Angeles, and New York, the nation's largest metropolitan areas, in addition to other areas. Chicago, Los Angeles, and New York are considered "self-representing" because their populations are over 2.5 million. Prices from self-representing areas are weighted by their actual populations. Other areas with populations below 2.5 million are "non-self-representing." Prices collected from non-self-representing areas are used to represent all areas of similar size in their Census region. Indices are constructed for each area, with items weighted by their relative importance to that population. Those area-specific indices are used to construct national indices. Within each area, a sample of outlets is selected to represent local and regional markets.

Outlets are assigned several “entry level items” selected from “major groups” for price data collection. “Food and beverages” is a major group, and includes both FAH and food away from home. An example of an entry level item in the food and beverages major group would be “bread,” which “includes all varieties of white bread and bread other than white” and “may be sold fresh, day old, frozen, or refrigerated.” Data collectors use a standardized protocol for probability selection of items that is described on the BLS website.

A stated goal of the FAH and other CPIs is that they reflect constant quality. Given regular changes in market offerings, the CPI survey must adjust for discontinued and reformulated products, while maintaining a constant-quality measure. In these instances, the data collector will select an item that is similar to the old item and identify the new data point as a substitution. Substitutions and their prices are reviewed by commodity analysts. If the commodity analyst deems it appropriate, the new item may be quality adjusted. Quality adjustment procedures are described in detail on the BLS CPI website.

The BLS is a principal statistical agency with a mission to provide information for statistical, and not commercial, purposes. CPI survey data is collected with strict confidentiality conditions. Raw data is not released to the public and processed data is embargoed for a period before release so that unauthorized users cannot access it before the public is able to do so.

Differences between the F-MAP to the CPI

The BLS FAH CPI differs from the F-MAP in several important ways (table 1). First, the underlying source data are different. In particular, the source data for the FAH CPI program are based on stratified probability surveys (i.e., CPI survey and the Consumer Expenditure Survey) whereas the source data for F-MAP are based on a census of retailers that agree to participate in

IRI marketing data collection. Although projection factors are used to weight the IRI scanner data to be representative of the store population in the United States, selection bias may be an issue. More importantly, the number of products in the IRI retail data far surpasses that of the CPI, which allows for more variability in the price information collected by the IRI retail scanner dataset.

The food categories and geographic area coverage of the two price index series also differ. In addition to a national-level FAH price index, the BLS publishes national-level price indices for six major categories (cereals and bakery products; meats, poultry, fish and seafood, and eggs; dairy and related products; fruits and vegetables; nonalcoholic beverages; and other FAH) and over 50 disaggregated food items under these broad FAH categories. At the Census Region, Census Division, and for 23 select metropolitan areas, the BLS also produces a price index for FAH, and for more recent years (2018 onward), price indices for these six broad FAH categories.

The F-MAP includes price indices for 90 food categories for the four Census Regions and in 10 select metropolitan areas that overlap geographically with the BLS 23 metropolitan areas. The F-MAP food categories can be aggregated to align with the six broad FAH categories from the CPI, but the detailed categories from each dataset differ considerably.

The F-MAP and the CPI are also conceptually different. Notably, the CPI can be used to compare prices over time but not across areas since the composition of the basket of goods vary substantially across areas (BLS 2023). In contrast, the F-MAP is constructed to compare prices over time and across areas. In addition, the CPI provides a reasonable estimate for price changes in urban areas whereas the F-MAP covers both urban and rural areas. Another difference between the CPI and F-MAP is the formula. The CPI is constructed in two steps. In the first step, the lower-level price indices are constructed using a modified geometric mean formula:

(1) $$cpi^{i,a}_{m,m-1} = \prod_{j\in(i,a)} \left[{}^{p_{j,m}}\!/_{p_{j,m-1}}\right]^{\left({}^{w_{j,0}}\!/_{\sum_{j\in(i,a)} w_{j,0}}\right)},$$

where $cpi^{i,a}_{t,t-1}$ is the lower-level price index for item i in area a between months m and m-1, $p_{j,m}$ and $p_{j,m-1}$ are the individual prices for item-area $j \in (i, a)$ for m and m—1, respectively, and $w_{j,0}$ are the expenditure shares for j in base period 0. In the second step, these lower-level price indices are aggregated into broad category, I, in broad area, A, using a modified Laspeyres index:

(2) $$CPI^{I,A}_{0',m} = CPI^{I,A}_{0',v} \times \frac{\sum_{i\in I,a\in A} W^{v+1}_{ia} \times cpi^{i,a}_{0,m}}{\sum_{i\in I,a\in A} W^{v+1}_{ia} \times cpi^{ia}_{0,v}},$$

where $CPI^{I,A}_{0',v}$ is the aggregate price index of price change from base period of the aggregate index, $0'$ to pivot month v (i.e., month prior to when aggregation weights are updated, usually December), $cpi^{i,a}_{0,m}$ and $cpi^{i,a}_{0,v}$ are the lower-level index of price change from base period 0 to period m or v for disaggregated item $i \in I$ and disaggregated area $a \in A$, and W^{v+1}_{ia} are the aggregation weights (BLS 2023). Both the lower-level geometric mean and upper level Laspeyres are bilateral index formulas. The F-MAP also includes four different bilateral price indices including the Laspeyres but also includes two multilateral indices, the GEKS and CCD, which are better suited for making price comparisons across geographies.

Quantifying the Effects of Data and Formula on Differences Between the F-MAP and CPI

To compare the two sets of price indices, we quantify how the underlying source data and price index formulas contribute to differences between the F-MAP and CPI. First, we construct price indices using the same underlying source data as the F-MAP (i.e., IRI Infoscan) but with the same conceptual and methodological approach as the CPI (equations 1 and 2). We call these Geometric Mean-Laspeyres (GML) scanner price indices. This will allow us to measure the effect of the underlying source data on the differences between CPI and the F-MAP. The GML price indexes are constructed for national and regional price indices for an aggregate FAH category and six food subcategories that closely align with the CPI major food categories. Appendix table 1 shows how the 90 EFPGs used to construct price indices in the F-MAP correspond to the 6 major CPI FAH subcategories.

We then quantify the differences between the two price measures (CPI versus F-MAP Laspeyres) in a series of linear regressions. In the first linear regression, we estimate the amount of variation in the F-MAP for each item (i.e., FAH; cereals and bakery products; protein foods; dairy and related products; fruits and vegetables; nonalcoholic beverages; other FAH), *I*, and region (i.e., nation and Census region), *R*, that can be explained by its corresponding CPI:

(1) $$F\text{-}MAP^{I,R} = \alpha^1 + \beta^1 CPI^{I,R} + \varepsilon^{I,R},$$

where $\varepsilon^{I,R}$ is the residual. The residual is the variation in the F-MAP that cannot be explained by the CPI for a particular *I*-*R*, and is used as the dependent variable in the second linear regression:

(2) $$\varepsilon^{I,R} = \alpha^2 + \beta^2 GML^{I,R} + u^{I,R},$$

where $u^{I,R}$ is the residual. The adjusted R^2 from the second step is the percentage of variation not explained by the CPI that can be explained by differences in data sources between the F-MAP and the CPI. This is called the data effect. One less the adjusted R^2 is the percentage of variation not explained by the CPI that can be attributed to other differences including formula effects.

From the first regression (Table 2a), roughly 4 percent of the variation in the national FAH F-MAP Laspeyres can be explained by the CPI leaving 96 percent of the unexplained variation attributable to differences in underlying source data and the formula between the two price measures. This results in a low correlation coefficient between the two FAH price measures of about 0.2 (=√0.04).

Table 2a here

The correlation between the F-MAP Laspeyres and CPI vary across regions and FAH subcategories. The FAH F-MAP Laspeyres for the Midwest, West and South are more highly correlated with their CPI counterparts with correlation coefficients ranging between 0.30 and 0.42 compared to that in the Northeast (0.09). The F-MAP Laspeyres for protein foods (i.e., beef, pork, poultry and eggs) is the most correlated with the CPI for protein foods (0.92) and the least correlated with cereals (0.16) and other foods (e.g., snacks, ready-to-eat and ready-to-heat meals, sugars and sweeteners, candy, fats and oils) (0.30). Overall, this shows that the relatively low correlation between the F-MAP Laspeyres at the national level for FAH and the CPI may be driven in part by differences in the FAH F-MAP Laspeyres in the Northeast and for FAH subcategories of cereals and bakery products and other foods.

At the national level, F-MAP Laspeyres understates inflation relative to CPI for FAH, as shown by the coefficient on the CPI being less than 1.0. In other words, a one-unit increase in the national CPI for FAH leads to less than a one-unit increase in the national F-MAP Laspeyres for FAH. However, this effect also varies by Census Region and FAH subcategory. F-MAP Laspeyres overstates inflation relative to CPI in the Midwest, South, and West regions and for the fruits and vegetables and beverages subcategories.

Based on the adjusted R^2 from the second linear regression (table 2b), differences in the underlying source data explain the majority of differences between the F-MAP and CPI. For FAH at the national level, the adjusted R^2 is 0.79, meaning 79% of the unexplained variation from the first regression can be explained by the GML scanner price index. And since the only difference between the GML scanner price index and the CPI is the underlying source data, then the remaining 21% of unexplained variation between F-MAP Laspeyres and CPI is attributable to the differences in formula.

Table 2b here

Similar to the first regressions, this varies across regions and FAH subcategories. The data effect dominates for cereals and bakery products (adjusted $R^2 = 0.89$), fruits and vegetables (0.56) and other foods (0.91) as well as for the Northeast (0.91), Midwest (0.72) and the South (0.69). This means that the differences in prices and sales collected by IRI Infoscan compared to the prices and household expenditures collected by BLS explain the majority of the differences in variation between the two price measures for most of the Census regions and half of the FAH subcategories.

Conclusion and Discussion

The CPI is a premier measure of price inflation of the U.S. economy. It is valuable for its timeliness, coverage of sectors across economy, extensive historical data, and rigorous and transparent methodology but has limited use for food and nutrition policy analysis. Current food and nutrition analysis require price information for more, detailed foods and geographic areas not covered in the CPI. The ERS created the F-MAP to fill this gap, which differs importantly from the CPI along several dimensions.

The underlying source data and price index formulas cause these price measures to differ in important ways. The F-MAP uses retail scanner data, which is a census of stores that agree to release data to ERS and have greater than $2 million in sales, in conjunction with a number of bilateral and multilateral price index formulas to measure price variation over geographic areas and time. In contrast, the CPI uses a stratified sample of stores for price collection and households for expenditures to construct a geometric mean-Laspeyres price index to measure price variation over time.

We quantify the contribution of differences in data and formula between the two price measures in a number of steps. First, we calculate price indices using the same formula as the CPI (geometric mean-Laspeyres price index formula) and the same source data as the F-MAP (retail scanner data) for FAH, a number of FAH subcategories, the nation and Census regions, referring to them as GML scanner data price indices. We then calculate the total amount of variation in the F-MAP across these foods and regions that cannot be explained by the CPI. Lastly, using the GML scanner data price indices, we partition the unexplained variation between the F-MAP and CPI into the data and formula effects.

The correlation between the F-MAP and CPI for national FAH seems low (0.2) but varies across regions and FAH subcategories. The F-MAP for FAH for the Midwest, South and West Census regions are moderately correlated with their CPI counterparts whereas the F-MAP for FAH for the Northeast is comparatively low (0.09). According to Levin et al. (2018), InfoScan store count coverage tended to be very low in parts of the Midwest and the Northeast, which may account for the low correlation between the F-MAP for the Northeast with the CPI for the Northeast. Even though the Infoscan stores are weighted to be representative of sales and counts of total food stores in the United States, it is only representative for grocery stores with more than $2 million in sales, excluding smaller, independent food stores. Indeed, the U.S. Census Bureau's County Business Patterns shows that the Northeast has a higher percentage of retail stores (51%) that employee less than 5 employees whereas the other Census regions have comparatively lower proportion of stores with less than 5 employees (41%-45%). The same degree of variability is also apparent when examining the FAH subcategories. The other foods subcategory could show low correlation between F-MAP and IRI due to substantial differences in the number and type of products included in the underlying data sources, but it is less clear as to why the cereals and bakery product F-MAP is less correlated with its CPI counterpart.

The differences in underlying data sets between the two price measures explains the majority of the difference between the CPI and F-MAP at the national level, for 3 out of 4 Census regions, and 3 out of the 6 FAH subcategories. The IRI Infoscan contains prices and sales for store that agree to release their data to ERS (nonprobability sample) whereas both the prices and household expenditures used for constructing the CPI price index weights are from separate stratified probability samples. Although the price and sales information in the IRI Infoscan are adjusted with post-stratification weights to make the IRI sample look more like total stores and sales for the United States, it is unclear whether these stores are representative of the variety or product assortment of excluded stores. The coverage of the IRI Infoscan also favors larger stores, which may price differently than smaller stores. The prices collected are also different across the two data sets: BLS collects the regular prices on foods and beverages whereas the prices in the IRI Infoscan are the price sold less discounts from sales. The latter is probably a better representation of prices consumers face while making purchases. Lastly, the IRI Infoscan database is considerably larger than that of the prices collected by BLS, both in number of retailers and number of price observations, allowing for more variability in the price information used in the F-MAP.

There are limitations to this research that we plan to explore in future iterations. The current research examines short-run relationships between the CPI, GML scanner data price indices, and

F-MAP-Laspeyres price indices but long-run relationships may also exist. As a next step, we plan to extend this analysis to 2016-20 and analyze longer-run relationships between the different indices using cointegration analysis.

References

Afshin A., Peñalvo J.L., Del Gobbo L., Silva J., Michaelson M., et al., 2017. The Prospective Impact of Food Pricing on Improving Dietary Consumption: A Systematic Review and Meta-Analysis. *PLOS ONE* **12**(3): e0172277.

Andreyeva, T., Long, M.W., and Brownell, K.D., 2010. The Impact of Food Prices on Consumption: A Systematic Review of Research on the Price Elasticity of Demand for Food. *American Journal of Public Health* **100**: 216-222.

Bureau of Labor Statistics. 2023. Handbook of Methods: Consumer Price Index. Available at https://www.bls.gov/opub/hom/cpi/home.htm. Accessed on 18 April 2023.

Çakır, M., Beatty, T.K., Boland, M.A., Park, T.A., Snyder, S. and Wang, Y., 2018. Spatial and Temporal Variation in the Value of the Women, Infants, and Children Program's Fruit and Vegetable Voucher. *American Journal of Agricultural Economics* **100**(3): 691-706.

Christensen, G. and Bronchetti, E.T., 2020. Local Food Prices and the Purchasing Power of SNAP benefits, *Food Policy*, 95(C).

Deaton, A. S., 1988. Quality, quantity, and spatial variation of price. *American Economic Review,* 78: 418–430.

Gregory, C.A. and Coleman-Jensen, A., 2013. Do High Food Prices Increase Food Insecurity in the United States? *Applied Economic Perspectives and Policy* **35**(4): 679-707.

Gundersen, C., Strayer, M., Dewey, A., Hake, M., and Engelhard, E., 2022. Map the Meal Gap 2022: An Analysis of County and Congressional District Food Insecurity and County Food Cost in the United States in 2020. Feeding America.

Handbury, J. and Weinstein, D. E., 2015. Goods prices and availability in cities. *The Review of Economic Studies*, 82(1): 258-296.

Levin, D., Noriega, D., Dicken, D., Okrent, A., Harding, M., & Lovenheim, M. (2018). *Examining Food Store Scanner Data: A Comparison of the IRI InfoScan Data With Other Data Sets, 2008-2012*, TB-1949. U.S. Department of Agriculture, Economic Research Service.

Muth, M. K., Kinney, S., Gargano, M., Looby, C., and Siegel, P., 2021. *User documentation: Store weights for InfoScan data*, 2012-2018. RTI International.

Muth, M.K., Karns, S. and Zhen, C., 2022, *Overview of Food Code Mapping for ERS Food Purchase Groups and the Monthly Food-at-Home Price Database*. RTI International. Available at https://www.rti.org/publication/overview-food-code-mapping-ers-food-purchase-groups-and-monthly-food-home-price-database/fulltext.pdf.

Muth, M. K., Sweitzer, M., Brown, D., Capogrossi, K. L., Karns, S. A., Levin, D., Okrent, A., Siegel, P., & Zhen, C. (2016). *Understanding IRI Household-Based and Store-Based Scanner Data*, TB-1942. U.S. Department of Agriculture, Economic Research Service.

Todd, J.E., Leibtag, E., and Penberthy, C., 2011. *Geographic Differences in the Relative Price of Healthy Foods*, EIB-78, U.S. Department of Agriculture, Economic Research Service.

Todd, J. E., Mancino, L., Leibtag, E., and Tripodo, C., 2010. *Methodology Behind the Quarterly Food-at-Home Price Database*, TB-1926, U.S. Department of Agriculture, Economic Research Service.

USDA Economic Research Service. 2023. Food Price Outlook. https://www.ers.usda.gov/data-products/food-price-outlook/. Accessed 4 April 2023.

U.S. Department of Health and Human Services and U.S. Department of Agriculture. *2015-2020 Dietary Guidelines for Americans,* 8th edition. Available at https://health.gov/sites/default/files/2019-09/2015-2020_Dietary_Guidelines.pdf. Accessed on 04 April 2023.

Zhen, C., Finkelstein, E.A., Karns, S.A., Leibtag, E.S. and Zhang, C., 2019. Scanner Data-Based Panel Price Indexes. *American Journal of Agricultural Economics* **101**(1): 311-329.

Table 1: Comparison of F-MAP with CPIs

	BLS: Food-at-home CPIs	ERS: F-MAP	ERS: Geometric Mean-Laspeyres (GML) scanner price indices
Sampling frame			
Prices	Commodities and services survey: multistage probability sample (~94,000 food and nonfood price quotes)	IRI Infoscan: census of stores that agree to release ERS information (~450,000 food UPCs per month)	IRI Infoscan: census of stores that agree to release ERS information (~450,000 food UPCs per month)
Consumption weights	Consumer Expenditure Survey: a multistage probability sample (~6,000 households per year)	IRI Infoscan: census of stores that agree to release ERS information (~450,000 food UPCs per month)	IRI Infoscan: census of stores that agree to release ERS information (~450,000 food UPCs per month)
Population	Urban areas across the United States	Urban and rural areas across the contiguous United States	Urban and rural areas across the contiguous United States
Methods			
Concept	Price comparisons over time	Price comparisons over time and areas	Price comparisons over time
Formula	Modified geometric mean formula for lower-level price indices and modified Laspeyres formula for aggregating lower-level price indices	Multilateral panel prices: rolling window Gini-Elteto and Koves-Szulc (GEKS) and Caves, Christensen and Diewert (CCD) formulas Bilateral panel prices: Laspeyres, Paasche, Törnqvist, and Fisher Ideal	Modified geometric mean formula for lower-level price indices and modified Laspeyres formula for aggregating lower-level price indices
Base period	Geographic area dependent, 1982-84 = 100[1], e.g., for Census regions, base period reflects prices in 1982-84 for a the Census region	National, 2016-18 = 100	Geographic area dependent, mostly 1982-84 = 100[1], e.g., for Census regions, base period reflects prices in 1982-84 for a the Census region

Table 1: Comparison of F-MAP with CPIs (continued)

	BLS: Food-at-home CPIs			ERS: F-MAP			ERS: Geometric Mean-Laspeyres (GML) scanner price indices		
Coverage	***National***	***Regional***	***SMA***	***National***	***Regional***	***SMA***	***National***	***Regional***	***SMA***
Foods	6 major food categories[5]; > 100 lower-level food prices within the major food categories	6 major food categories[5]	6 major food categories	90 foods and beverages	90 foods and beverages	90 foods and beverages	6 major FAH categories consistent with BLS major food categories[5]	6 major food categories consistent with BLS major food categories[5]	6 major food categories consistent with BLS major food categories[5]
Areas	na	4 Census regions[3]; 9 Census divisions[4]	23 metropolitan areas including those covered under F-MAP[2]	na	4 Census regions[3]	10 SMAs[2]	na	4 Census regions[3]	10 SMAs[2]

Time	Starting in 1935 for some major FAH categories and various years for others and lower-level food prices	Starting in 1977 for Census regions; starting in 2017 for Census divisions	Starting in 1953 for select metropolitan areas; starting in 2018 for 6 major food categories in SMAs	Currently available for 2016 to 2018 but to be updated for 2012-2020.	Currently available for 2016 to 2018 but to be updated for 2012-2020.	Currently available for 2016 to 2018 but to be updated for 2012-2020.	2016 to 2018	2016 to 2018	2016 to 2018

[1] For a handful of major FAH categories, namely in nonalcoholic beverages, the base period is December 1977 through November 1978.

[2] The SMA coved under F-MAP are: Atlanta-Sandy Springs-Roswell, Georgia; Boston-Cambridge-Newton, Massachusetts; Chicago-Naperville, Elgin, Illinois; Dallas-Fort Worth-Arlington, Texas; Detroit-Warren-Dearborn, Michigan; Houston-The Woodlands-Sugar Land, Texas; Los Angeles-Long Beach-Anaheim, California; Miami-Fort Lauderdale-West Palm Beach, Florida; New York City-Newark- Jersey City, New York and New Jersey; and Philadelphia-Camden-Wilmington, Pennsylvania and Delaware.

[3] The four Census regions are Northeast, South, Midwest and West.

[4] The nine Census divisions are: New England, Middle Atlantic, South Atlantic, East South Central, West South Central, East North Central, West North Central, Mountain and Pacific.

[5] The major FAH categories in the CPI are: cereals and bakery products; meats, poultry, fish and seafood, and eggs; dairy and related products; fruits and vegetables; nonalcoholic beverages; and other FAH.

Table 2a. Linear Regression of F-MAP Laspeyres Index on CPI

Explanatory variables	National FAH F-MAP	Northeast FAH F-MAP	Midwest FAH F-MAP	South FAH F-MAP	West FAH F-MAP	National Cereal F-MAP	National FV F-MAP	National Dairy F-MAP	National Meat F-MAP	National Beverage F-MAP	National Other Food F-MAP
	Dependent Variable										
Constant	0.273	1.204	-0.357	-0.463	-0.099	0.754	-0.217	0.252	0.128	-0.016	6.515
	(0.637)	(0.342)	(0.620)	(0.800)	(0.423)	(0.269)	(0.222)	(0.085)	(0.065)	(0.249)	(2.970)
CPI	0.753	-0.179	1.385	1.488	1.124	0.253	1.252	0.754	0.887	1.024	-5.462
	(0.637)	(0.342)	(0.620)	(0.800)	(0.423)	(0.269)	(0.222)	(0.085)	(0.065)	(0.249)	(2.970)
R^2	0.039	0.008	0.128	0.092	0.172	0.025	0.483	0.698	0.846	0.333	0.091
N	36	36	36	36	36	36	36	36	36	36	36

Table 2b. Linear Regression of OLS Residual from Table 2a on GML Scanner Price Index

Explanatory variables	National FAH Residual	Northeast FAH Residual	Midwest FAH Residual	South FAH Residual	West FAH Residual	National Cereal Residual	National FV Residual	National Dairy Residual	National Meat Residual	National Beverage Residual	National Other Food Residual
	Dependent Variable										
Constant	-2.019	-2.071	-1.510	-1.966	-1.348	-1.230	-0.524	-0.327	-0.162	-0.653	-1.980
	(0.176)	(0.110)	(0.160)	(0.227)	(0.265)	(0.073)	(0.080)	(0.078)	(0.064)	(0.140)	(0.109)
GML	2.011	2.060	1.507	1.959	1.342	1.228	0.512	0.326	0.160	0.652	1.988
	(0.175)	(0.109)	(0.160)	(0.226)	(0.264)	(0.073)	(0.079)	(0.078)	(0.063)	(0.140)	(0.110)
R^2	0.794	0.913	0.723	0.688	0.433	0.892	0.555	0.339	0.159	0.389	0.906
N	36	36	36	36	36	36	36	36	36	36	36

Notes: Standard errors are in parentheses.

Source: Authors' calculations based on F-MAP price indexes constructed from Circana (formerly IRI) Infoscan data and the CPI (Bureau of Labor Statistics 2023).

Appendix Table A1. Crosswalk Between ERS Food Product Groups (EFPGs) and BLS CPI Major FAH Categories

EFPG		CPI	
Code	*Description*	*Item code*[1]	*Description*
10000	Whole-grain breads	_SAF111	Cereals and bakery products
10025	Whole-grain rice and pasta	_SAF111	Cereals and bakery products
10050	Whole-grain breakfast grains	_SAF111	Cereals and bakery products
10075	Whole-grain flour, bread mixes, and frozen dough	_SAF111	Cereals and bakery products
15000	Non-whole-grain breads	_SAF111	Cereals and bakery products
15025	Non-whole-grain rice and pasta	_SAF111	Cereals and bakery products
15050	Non-whole-grain breakfast grains	_SAF111	Cereals and bakery products
15075	Non-whole-grain flour, bread mixes, and frozen dough	_SAF111	Cereals and bakery products
73030	Baked goods	_SAF111	Cereals and bakery products
73040	Cake and cookie mixes	_SAF111	Cereals and bakery products
74000	Whole-grain breakfast cereal	_SAF111	Cereals and bakery products
74050	All other breakfast cereal	_SAF111	Cereals and bakery products
20000	Fresh potatoes	_SAF113	Fruits and vegetables
20075	Canned potatoes	_SAF113	Fruits and vegetables
21500	Fresh other starchy vegetables	_SAF113	Fruits and vegetables
21525	Fresh-cut other starchy vegetables	_SAF113	Fruits and vegetables
21550	Frozen other starchy vegetables	_SAF113	Fruits and vegetables
21575	Canned other starchy vegetables	_SAF113	Fruits and vegetables
23000	Fresh tomatoes	_SAF113	Fruits and vegetables

23075	Canned tomatoes	_SAF113	Fruits and vegetables
24500	Fresh other red and orange vegetables	_SAF113	Fruits and vegetables
24525	Fresh-cut other red and orange vegetables	_SAF113	Fruits and vegetables
24550	Frozen other red and orange vegetables	_SAF113	Fruits and vegetables
24575	Canned other red and orange vegetables	_SAF113	Fruits and vegetables
26000	Fresh dark green vegetables	_SAF113	Fruits and vegetables
26525	Fresh-cut dark green vegetables	_SAF113	Fruits and vegetables
26550	Frozen dark green vegetables	_SAF113	Fruits and vegetables
26575	Canned dark green vegetables	_SAF113	Fruits and vegetables
27500	Fresh/dried beans, lentils, and peas	_SAF113	Fruits and vegetables
27550	Frozen beans, lentils, and peas	_SAF113	Fruits and vegetables
27575	Canned beans, lentils, and peas	_SAF113	Fruits and vegetables
29000	Fresh other/mixed vegetables	_SAF113	Fruits and vegetables
29025	Fresh-cut other/mixed vegetables	_SAF113	Fruits and vegetables

EFPG		CPI	
	Description	*Item code*[1]	*Description*
29050	Frozen other/mixed vegetables	_SAF113	Fruits and vegetables
29075	Canned other/mixed vegetables	_SAF113	Fruits and vegetables
30000	Fresh whole fruit	_SAF113	Fruits and vegetables
30025	Fresh-cut fruit	_SAF113	Fruits and vegetables
30050	Frozen fruit	_SAF113	Fruits and vegetables
30075	Canned fruit	_SAF113	Fruits and vegetables
30090	Dried fruit	_SAF113	Fruits and vegetables
40000	Whole milk	_SEFG	Dairy and related products
40030	Whole-fat cream and sour cream	_SEFG	Dairy and related products
40060	Whole-milk yogurt	_SEFG	Dairy and related products
43000	Reduced-fat, low-fat, and skim milk	_SEFG	Dairy and related products
43030	Reduced-fat and low-fat cream and sour cream	_SEFG	Dairy and related products
43060	Reduced-fat, low-fat, and skim-milk yogurt	_SEFG	Dairy and related products
46000	Cheese and cream cheese	_SEFG	Dairy and related products
46050	Processed cheese	_SEFG	Dairy and related products
72020	Flavored milk and other sweetened milk-based beverages	_SEFG	Dairy and related products
73050	Ice cream and other milk-based desserts	_SEFG	Dairy and related products
50000	Fresh beef, pork, lamb, veal, and game	_SAF112	Meats, poultry, fish, and eggs
50050	Frozen beef, pork, lamb, veal, and game	_SAF112	Meats, poultry, fish, and eggs
50075	Canned beef, pork, lamb, veal, and game	_SAF112	Meats, poultry, fish, and eggs
51500	Fresh chicken, turkey, and game birds	_SAF112	Meats, poultry, fish, and eggs
51550	Frozen chicken, turkey, and game birds	_SAF112	Meats, poultry, fish, and eggs
51575	Canned chicken, turkey, and game birds	_SAF112	Meats, poultry, fish, and eggs
53000	Fresh fish and seafood	_SAF112	Meats, poultry, fish, and eggs
53050	Frozen fish and seafood	_SAF112	Meats, poultry, fish, and eggs

53075	Canned fish and seafood	_SAF112	Meats, poultry, fish, and eggs
56000	Bacon, sausage, and lunchmeats	_SAF112	Meats, poultry, fish, and eggs
57500	Egg and egg substitutes	_SAF112	Meats, poultry, fish, and eggs
35000	Fresh 100% fruit and vegetable juices	_SAF111	Nonalcoholic beverages
35050	Frozen 100% fruit and vegetable juices	_SAF112	Nonalcoholic beverages
35075	Canned/shelf-stable 100% fruit and vegetable juices	_SAF113	Nonalcoholic beverages
72000	Sweetened coffee and tea	_SAF114	Nonalcoholic beverages
72010	Unsweetened coffee and tea	_SAF114	Nonalcoholic beverages
72030	Low-calorie beverages	_SAF114	Nonalcoholic beverages
72040	All other caloric beverages	_SAF114	Nonalcoholic beverages
72060	Water	_SAF114	Nonalcoholic beverages

	EFPG	CPI	
Code	*Description*	*Item code*[1]	*Description*
72050	Alcohol	_SAF116	Alcoholic beverages, at home
54500	Nuts and seeds	_SAF115	Other food at home
54550	Nut and seed butters and spreads	_SAF115	Other food at home
73000	Sweeteners	_SAF115	Other food at home
73010	Jellies/jams	_SAF115	Other food at home
73020	Candy	_SAF115	Other food at home
73060	All other desserts	_SAF115	Other food at home
75000	Whole-grain savory snacks	_SAF115	Other food at home
75050	All other savory snacks	_SAF115	Other food at home
76000	Vitamins and meal supplements	_SAF115	Other food at home
77000	Baby food	_SAF115	Other food at home
78000	Infant formula	_SAF115	Other food at home
59000	Tofu and meat substitutes	_SAF115	Other food at home
60000	Ready-to-eat foods	_SAF115	Other food at home
62500	Frozen/refrigerated ready-to-heat foods	_SAF115	Other food at home
65000	Shelf-stable ready-to-heat foods and soups	_SAF115	Other food at home
67500	Shelf-stable meal kits	_SAF115	Other food at home
70000	Fats and oils	_SAF115	Other food at home
70050	Salad dressing	_SAF115	Other food at home
71000	Condiments, gravies, and sauces	_SAF115	Other food at home
71050	Dry spices	_SAF115	Other food at home
99999	Not coded	na	na

[1]The BLS series ID consists of the 4 parts: (1) the first 2 digits indicate the database name (all urban consumers, urban wage earners and clerical workers, all urban consumers (chained CPI), and average price data); (2) the next two digits indicate whether the series is seasonally or not seasonally adjusted; (3) the next 4 digits indicate the geographic coverage; and (4) the last digits indicate the item. This column shows the last digits of the series ID and the underscore indicates that the item code is extension to a longer series ID.