

Data collection sources and methods:

www.globalproductprices.com

This dataset became possible because of the widespread move to online shopping during the pandemic. In many countries, the largest grocery stores developed online presence where prices are essentially the same as in their physical locations. It became possible therefore to collect data on food prices from large, established grocers by looking at their online stores.

A “benchmark” price

We identified: 1) the largest grocery stores in each country with significant online presence and 2) the most established brands of each product category in that country. The selection of the products aims at popular mid-level priced items. Moreover, the products have clear characteristics that make them comparable across countries. As such, these prices are a “benchmark” of the prices charged by the largest brands in the largest stores of each country.

Price variation across stores

The prices of large retailers are called “benchmarks” as other grocers have to take them into account when they set prices. Smaller competitors cannot charge more than the market leader on a continuous basis as they lose sales. They also don’t have an incentive to charge much less than the market leader as they lose revenue and, often, they *cannot* charge much less as the market leader employs economies of scale to bring down prices. In other words, the largest chains are price setters while the smaller stores adjust their pricing to them. As a result, the prices of the large retailers are indicative of the prices in general.

You can, for example, compare the prices that we collect from Walmart in the U.S. which is the largest grocer in the country ([here](#)) with the official government data from the Bureau of Labor Statistics ([BLS](#)). It is clear that the national average prices even in this large and diverse country are very close to the prices of the market leader.

Regional price variation

There may be regional price differences in parts of the country where the large stores do not have presence, but these are typically more remote rural areas where the population is relatively sparse. Rapid urbanization in recent decades has brought more and more people into large metropolitan areas. For example, over 77 percent of people in advanced countries live in urban settings. In the U.S. that is over 82 percent. The conglomeration has also allowed the establishment of large stores in close proximity to where most people live.

From that perspective, while urban vs rural price differences may exist, urban prices are in some sense more *representative* of the prices paid by most people. Not surprisingly, the most widely used and cited measure of price inflation - the Consumer Price Index - is also based on the consumption patterns of urban consumers.

Why not use official government data

Unfortunately, relatively few governments publish regular up-to-date information on food prices. While they regularly publish a food price inflation index, they do not make available the underlying price level data. In addition, even when they publish such information, the products and the timing of data releases differ by country. Often the product categories are very general. For example, the categories rice, beef, and flour are not revealing of the exact products included in the category. That complicates cross country comparison and limits the scope of the data set to a handful of countries.

As an alternative, we collect more targeted data with clearly identified sources that can be compared across countries and over time. Of course, where such data are available, even with a lag, we use the official statistics to make sure we are not off the mark.

Why not use online surveys and user input

Online input by locals might have been the only way to collect these data until a few years ago when online food shopping was not widespread. It provides valuable information, but one

cannot be sure when the data were collected, from how many respondents, and who these respondents are.

With our method, we are transparent about where the data come from. The approach also allows very timely updates which is valuable as food prices, along with energy prices, are the most volatile components of the CPI.

How we collect the data

The sources are identified and continuously verified by our in-house team. The data are also collected by our in-house team without scraping or automation. The reason is that we can preempt data mistakes stemming from a switch of a product location on the grocers' websites, special sale prices, change in measuring units and so forth. An automated system cannot (yet) run those checks to a satisfactory level of confidence.

Please address any questions to neven.valev@theglobaleconomy.com.