Information on the change in the calculation methodology of the consumer price index for food and non-alcoholic beverages

Until the end of 2023, in the field of price statistics, the data sources used to compile the consumer price index for food and non-alcoholic beverages were obtained from data collection in the field. Data collection was carried out directly in shops throughout the Slovak Republic, where residents usually do the shopping, and the prices that were surveyed were the so-called counter prices. Prices were determined at a some point in time during the first 20 days of the monitored month.

In order to improve and modernize statistical production, the Statistical Office of the Slovak Republic established cooperation with retail chains to implement transactional data on retail sales into the production of price statistics. This data source was analysed and its use in price statistics was consistently tested. From January 2024, the Statistical Office of the Slovak Republic moved from the experimental phase of creating statistical outputs to the official production process, when calculating the consumer price index by using transaction data for food and non-alcoholic beverages,.

Transaction data, also called scanner data, is data that retailers record when consumers make purchases by scanning barcodes. These data include, for each item sold in this period, the quantity sold and the selling/realization price at the level of item code.

The Statistical Office of the Slovak Republic has an access to aggregate values in the form of turnover and sold quantity at the level of individual sold products for each calendar week. The transaction data thus allows compiling an index from all the retailer's transactions, while the realization prices of products enter into the calculation and allows including many more items in the consumer price index compared to the traditional collection of prices.

For example, we can state that by the end of 2023, about 16,000 monitored product prices were surveyed in retail chains for 146 selected representatives. Currently, the calculation includes the complete range of sold products, which amounts to approx. 60,000 average product prices for food and non-alcoholic beverages.

The use of transaction data, in addition to other significant methodological changes, also implies a change in the pricing concept. Prices collected in the traditional way are replaced by average prices per unit of goods, which more accurately reflect the prices paid by consumers during the entire monitored period and take into account discounts and their impact on the amount of goods sold. Thus, the price surveyed at a specific time is not included in the calculation of price indices, but the average price per unit of goods for the monitored period determined as follows:

average price per product unit = turnover per product / number of units sold

Currently, the Statistical Office of the Slovak Republic collects transaction data for food and non-alcoholic beverages from the five largest retail chains on a weekly basis, which represents approximately 80% of the total turnover in retail trade for this area.

Data files contain turnover and sold amounts of individual goods in the form of a weekly aggregate (from Monday to Sunday). Structural validation of the file is conducted on the input to the information system. Subsequently, individual products are classified into the ECOICOP classification of individual consumption, which divides the basket of goods and services into divisions (2-digit), groups (3-digit), classes (4-digit) and subclasses (5-digit). For 01 – Food and non-alcoholic beverages ECOICOP classification, a national, more detailed 6-digit level of classification - ECOICOP6 - was defined to process data from scanners, which is common to the data of all retail chains currently cooperating with the Statistical Office of the Slovak Republic. The 6-digit level was defined in order to create homogeneous groups of products.

Each product enters the index calculation based on its importance. Therefore, filters are applied to the data that exclude some products from the calculation. Filtering removes products with an extreme change in price compared to the previous period, products of final sale (a drop in price and a significant decline in turnover at the same time) and low selling products.

After many analyses and following discussions within the European statistical system, the Statistical Office of the Slovak Republic decided to follow the path of a dynamic instead of a static approach. The dynamic approach better takes into account the currently sold products and is simpler from the point of view of the automation process. Products that are on sale simultaneously in both consecutive periods/months (t and t + 1, t + 1 and t + 2, t + 2 and t + 3, etc.) with turnover above a certain threshold are automatically selected. Each month, the set of individual products that enter the index is selected anew. At the lowest level of aggregation, the index is calculated based on a set of paired representative commodity items that are actually sold in two consecutive periods.

To calculate the consumer price index, the Jevonson formula is used at the lowest level of aggregation, which can be interpreted as the geometric mean of the price ratio:

$$I_{Jevons}^{m-1,m} = \prod_{i \in N} \left(\frac{p_i^m}{p_i^{m-1}} \right)^{\frac{1}{N}}$$

where *m* is the current period, *m-1* is the previous period and p_i is the price of the *i*-th product in the period.

From an axiomatic point of view, the Jevons index is clearly the index with the best properties. The axiomatic approach means that the index satisfies certain specific axioms or tests on the basis of which an appropriate index can be selected. The most preferred tests are proportionality test, symmetry test, time reversal test and transitivity test.

To calculate the price change at a higher level of aggregation of the ECOICOP classification, a Laspeyres-type index with fixed weights is used, referring to the previous year as the reference period (y-1):

$$P_A^{y,m/y-1} = \frac{\sum_{a \in A} w_a^{y-1} I_a^{y,m/y-1}}{\sum_{a \in A} w_a^{y-1}}$$

where $w_a^{y-1}$are the weights based on annual expenditure for all items. Subsequently, the indexes calculated in this way are chained to the base period. The use of this type of index arises from the Eurostat regulations.

The impact of the methodological change on the overall consumer price index expressed by the average absolute difference was 0.2 percentage points and on the consumer price index for food and non-alcoholic beverages by 0.9 percentage points.

Change to a new data source means that we joined the 12 countries in the European Union that use this data source when calculating inflation. In other countries, preparatory work is currently underway for the use of this data source.

The Statistical Office of the Slovak Republic will continue systematically extending the range of products, for which transaction data will be the source for calculating the consumer price index.