Article

Tracking the price of the lowest-cost grocery items, UK, experimental analysis: April 2021 to September 2022

How the prices of the lowest-cost products for 30 everyday items have changed since April 2021.

Contact:
Emily Hopson
cpi@ons.gov.uk
+44 1633 455 592

Release date:
25 October 2022

Next release:
To be announced

Table of contents

1 . Main points

- Highly experimental research, based on web-scraped supermarket data for 30 everyday grocery items, shows that the lowest-priced items have increased in cost by around 17% over the 12 months to September 2022: an increase from 7% over the 12 months to April 2022.

- Although not directly comparable, the rise in prices for the lowest-cost grocery items is similar to the 15% rise in the official measure of inflation for food and drink.

- There is considerable variation across the 30 items, with the prices for four items falling between September 2021 and September 2022, but the prices of 15 items rising by 15% or more.

- Vegetable oil showed the largest percentage increase and average price increase between April 2022 and September 2022, increasing 46% (80 pence per litre) during the period; in contrast, orange juice showed the largest decrease, decreasing 8% (6 pence per litre).

- The data provided in this article remain highly experimental; we have updated the methodology used for the choice of substitutes for missing products, meaning estimates will differ from the previous publication of this data: more information on why we have chosen this approach is available in Section 9: Strengths and limitations.

Data and analysis in this article have been produced using new, innovative methods and as a result are less robust than official statistics.

Please note, we have updated the methodology used for the choice of substitutes for missing products, meaning estimates will differ from the previous publication of these data: more information on why we have chosen this approach is available in Section 9: Strengths and limitations.

2 . Price changes for lowest-cost items

Using innovative analytical methods to track the lowest-priced grocery items

With rising prices seen across many goods and services, the Office for National Statistics (ONS) has looked at the question of how prices of everyday grocery items have changed for the lowest-cost products.

To try to answer this, we have updated previously published analysis, where we applied new and highly experimental methods, making use of web-scraped supermarket data to capture the price changes of everyday grocery items.

From April 2021 to September 2022, online grocery price quotes were collected from seven major supermarket retailers' websites. Prices were assessed for 30 everyday food and drink items, covering fresh fruit and vegetables, cupboard staples, chilled products, as well as meat and fish.

3 . Price movements for each grocery item

The lowest-priced everyday grocery items have seen a notable variation in price change, with some items showing increases of well over 20% from September 2021, while other items fell in price

For half of the 30 sampled items monitored (15 items), the average lowest price, across the retailers, increased at a faster rate than the latest available official consumer price inflation measure for food and non-alcoholic beverages (a 15% increase between September 2021 and September 2022). Caution should be taken when comparing with the official measure of consumer price inflation for food and non-alcoholic beverages, as it contains many more than the 30 items used in this analysis as well as using different aggregation methodology.

For nine items, the lowest-cost price increased by more than 20% since September 2021, and for three of those nine items the lowest-cost price rose by 40% or more. The items where the lowest prices rose at the fastest rate between September 2021 and September 2022 were:

- Vegetable oil (65%)

- Pasta (60%)

- Tea (46%)

For four of the 30 items, the lowest prices fell on average during the same period. The largest price decrease measured was for fruit orange juice with an 9% fall in price, followed by beef mince which saw an 7% decrease.

It is important to consider that for each of the 30 items, the overall figure can be made up of different price movements at the product level. This includes when we are missing product-level price data, for a retailer, the cheapest product could be replaced with a more expensive item for a number of months before returning to track the original lowest priced item.

Another example is vegetable oil (including sunflower oil and rapeseed oil) where prices rose by 65%. However, for vegetable oil, there are instances where we are missing some product-level price data, at some retailers, for the cheapest oils. In one case, this caused us to track the price of a more expensive oil. For these reasons, the estimate of 65% should be treated with additional caution.

We provide further explanation on issues with the data used for this analysis in Section 9: Strengths and limitations.

Figure 1: There was a substantial range of price movement for the lowest prices

Lowest price of selected 30 everyday groceries, item-level price changes, September 2022 compared with September 2021

Download the data

.xlsx

Table 1 shows over the 12 months to September 2022, vegetable oil, chips and milk were the top largest average price rises, in cash terms. However, beef mince shows the largest reduction, dropping 15 pence between September 2021 and September 2022.

Between April 2022 and September 2022, the three largest average price rises in cash terms were:

- Vegetable oil (up 80 pence for 1 litre to £2.58)

- Chips (up 27 pence to £1.37 for 1.5kg)

- Milk (up 25 pence to £1.52 for 4 pints)

In contrast, the largest average fall in the lowest averages prices measured between April 2022 and September 2022 were:

- Fruit orange juice (down 6 pence to 76 pence for 1 litre)

- Beef mince (down 5 pence to £1.95 pence for 500g)

Table 1: Average price of lowest-priced grocery items, September 2021 to September 2022 in pence

Item	Average size (units)	Price in September 2021	Price in April 2022	Price in September 2022	Difference between September 2021 and September 2022	Difference between April and September 2022
Vegetable oil	1000ml	156	178	258	102	80
Chips	1500g	99	110	137	38	27
Milk	2272ml (4 pints)	117	127	152	35	25
Chicken breast	600g	330	350	363	33	13
Tea	125g	67	97	97	30	0
Instant coffee	100g	131	135	155	24	20
Mixed frozen vegetables	1000g	76	89	100	24	11
Pasta	500g	38	54	61	23	7
Bread	800g	48	54	66	18	12
Crisps	150g	73	84	91	18	7
Ham	200g	160	165	176	16	11
Sausages	454g	89	92	105	16	13
Apples	6 pack	83	92	97	14	5
Biscuits	400g	40	45	54	14	9
Onions	1000g	63	67	75	12	8
Potato	2500g	94	96	106	12	10
Fish Fingers	250g	76	77	87	11	10
Cheese	255g	89	88	98	9	10
Tomatoes	250g	48	55	57	9	2
Breakfast cereal	500g	69	76	76	7	0
Tomato ketchup	460g	58	61	65	7	4
Bananas	5 pack	73	75	78	5	3
Baked beans	400g	30	31	34	4	3
Fruit squash	750ml	48	52	51	3	-1
Pizza	300g	84	86	87	3	1
Yoghurt	500g	55	55	58	3	3
Granulated sugar	1000g	66	67	66	0	-1

Rice	1000g	89	89	89	0	0
Fruit juice orange	1000ml	83	82	76	-7	-6
Beef mince	500g	210	200	195	-15	-5

Source: Office for National Statistics – Tracking the lowest cost grocery items

Notes

1. All price and difference columns are given in pence.

Figure 2 shows the extent to which prices have changed since April 2021. We have presented item indices with prices in April 2021 given a reference figure of 100. If an item was £2 in April 2021 and rose to £2.20, the item index would increase from 100 in April 2021 to 110, reflecting a 10% price increase. Alternatively, if the same item had fallen in price to £1.80, the index value would fall to 90 reflecting a 10% price reduction.

For the majority of items, over the year to September 2022 - most notably pasta, vegetable oil, and bread - prices for the lowest-priced everyday grocery items have increased.

However, the exact timing of price increases varied depending on the individual item. For example, the lowest price of pasta rose by 18% over October and November 2021, while vegetable oil prices rose 12% over June and July 2022.

Other items saw a more gradual increase or decrease in the lowest price. For example, the lowest price of orange fruit juice trended downward from November 2021 and was 9% lower in September 2022 compared with a year ago.

Several items had a very stable lowest price throughout the entire period, such as yoghurt and pizza.

Figure 2: Lowest-cost prices saw varied movements from April 2021 to September 2022

Lowest price of groceries, item-level price index, April 2021=100

Download the data

.xlsx

4 . Lowest price of 30 everyday items

Combining the lowest-cost items into an index shows that, overall, the prices of the cheapest items has risen since April 2021 broadly in line with official measures of inflation

Figure 3 shows that an overall groceries index that combines the lowest prices of 30 everyday items follows a broadly similar trend to official measures of inflation for the food and non-alcoholic beverages component of the consumer price index including owner occupiers' housing costs (CPIH). However, caution should be taken when comparing with the official measure of consumer price inflation for food and non-alcoholic beverages, as it contains many more than the 30 items used in this analysis as well as using different aggregation methodology.

While there is a lot of variation at the individual item level, overall the lowest prices of the 30 everyday items, weighted by retailer and item, rose by 17% in the year to September 2022. This includes a 9% increase since our previous analysis in April 2022.

Figure 3: Prices for the lowest-cost grocery items follow a broadly similar trend to the equivalent official measure of inflation for food and drink

Lowest cost grocery items, April 2021=100

Figure 3: Prices for the lowest-cost grocery items follow a broadly similar trend to the equivalent official measure of inflation for food and drink

Lowest cost grocery items, April 2021=100
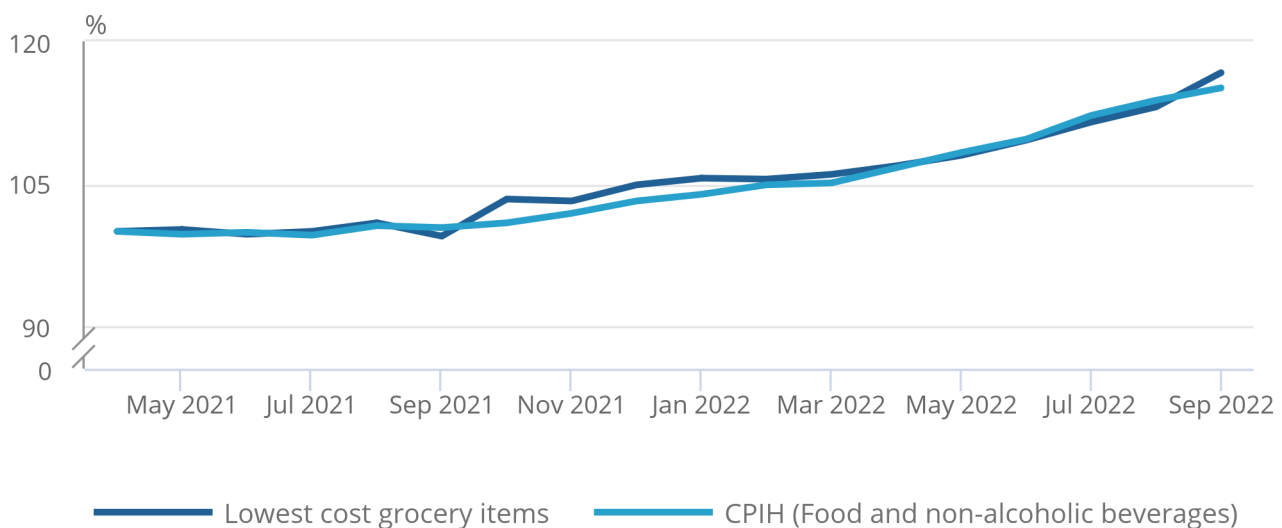
Source: Office for National Statistics – Tracking the lowest-cost grocery items

5 . Using web-scraped data

All data presented in this article is based on web-scraped supermarket data for 30 everyday grocery items. One limitation of this approach is that items may not always be available instore or online, which is reflected in the data collected by our web scrapers. As a result, the measure of the lowest price presented in this analysis can be sensitive to product availability and the specific products that are being substituted.

Different approaches to substituting items can result in very different trends. For this release, we have updated the methodology used for the choice of substitutes for missing products, meaning estimates will differ from the previous publication of this data. Instead of a fixed threshold, the 20% threshold for substitution has now been calculated on a rolling basis using the minimum and maximum prices in the current and previous month.

A drawback of both these approaches is that it can cause us to miss the entry of cheaper products for an item where the new products come in at a price point far enough below the current product. Section 9: Strengths and limitations explains the full limitations of this substitution method.

Value ranges often represent a substantial saving and, where they are not available, the price difference to the next lowest-priced available item is often large. Over the last year, there is evidence to suggest that there has been a growing number of retailers introducing lower-priced ranges online, which previously might not have been available online. Even if these ranges were in store, the data for those prices would not be captured in this article. This may affect changes in prices measured over the past year, but our data suggests this is less likely to affect price movements over the past six months.

6 . Lowest-cost grocery item data

Analysis of lowest-cost items, UK
Dataset | Released 25 October 2022
Data tables containing the item list and volumes, price change and indices published alongside the Office for National Statistics' (ONS') analysis of lowest-cost items.

7 . Glossary

Consumer price inflation

Consumer price inflation is the rate at which the prices of goods and services bought by households rise or fall. It is estimated by using price indices. For an overview of the indices and their uses, please see our article, Consumer price indices, a brief guide: 2017.

Consumer Prices Index including owner occupiers' housing costs (CPIH)

CPIH is the most comprehensive measure of inflation. It extends the Consumer Prices Index (CPI) to include a measure of the costs associated with owning, maintaining, and living in one's own home, known as owner occupiers' housing costs (OOH), along with Council Tax. Both are substantial expenses for many households and are not included in the CPI.

Web scraping

Web scraping is the activity or process of taking information from a website. During the coronavirus (COVID-19) pandemic, the Office for National Statistics (ONS) developed web-scraping capability as part of a previous project, which is explained in our Online price changes for high-demand products methodology. This work was expanded to cover a wider range of grocery products.

8 . Measuring the data

Data sources and quality

The data presented in this article is experimental. More information on experimental statistics is available in our guide to experimental statistics.

The web-scraped data have been collected from seven grocery retailers: Asda, the Co-op, Iceland, Morrisons, Sainsbury's, Tesco and Waitrose. Co-op data was not available and not included in the estimates from May 2022 onwards because of changes introduced to the Co-op website. However, this change showed minimal impact because of the small weight allocated to Co-op. More details on retailer weights are available in Section 8: Measuring the data.

For each item and retailer, the price of the lowest-priced product available was selected (after adjusting for the size of the product) over time to see how it changed.

These data differ from the data used to compile the official measures of inflation for food and non-alcoholic beverage items. Those measures comprise data for the 179 food and non-alcoholic drink items, with local price collectors collecting 50,000 prices by visiting sampled retailers in over 140 locations across the UK.

Item choice

The items have been chosen using data from the Department for Environment, Food and Rural Affairs (DEFRA) 2019/20 Family food datasets, which is a module within our Living Costs and Food Survey. The complete list of items was prepared in collaboration with external stakeholders.

These data enabled us to identify the grocery items most likely to be bought by lower income households.

A total of 30 items were chosen as a good trade-off between coverage of a high proportion of expenditure, and the costs of adding more items to the analysis.

The approach started with the items with the highest expenditure and the largest quantity bought by households in the lowest-equivalised income decile. Other factors such as the substitutability of items, the representativeness of comparable items and including a broad range of products were also factored in. For example, prices of minced beef are likely to respond to similar economic factors of other beef and meat products.

When selecting products, we only include products within a per item size band. For example, for sausages, we focus on products between 350g and 700g, inclusive. The size bands are used to define representative items and exclude more expensive bulk or large multipack products. This creates a more consistent comparison between retailers and addresses data quality issues, such as misclassifications and product sizing errors.

The size bands used are based on a manual review of common product sizing for the retailers, pragmatic decisions to maximise availability, and consistency of data.

Item labelling

To identify the basket items in the web-scraped data, we use manually defined keywords for each item. This builds on part of a previous Office for National Statistics (ONS) project, which is explained in our Online price changes for high-demand products methodology.

For example, to identify "apples" we might expect the product to be listed in the "Fresh Fruit" section of the retailer product catalogue and to contain the word "apples" in the product name. In practice, each item has an accompanying list of product catalogue sections, keywords to match in the product name, and a separate list of words to exclude products (for example, for "apples" we would not want to include "toffee apples").

This approach is simple to apply, and quick to replicate and add new items, but it does have some drawbacks.

Manual quality assurance is required to check that the keywords select the expected products. It is also hard to guarantee that all the target products have been found. Additionally, there is judgement required around which products to allow as representatives for an item, factoring in packaging, product composition, reliability of size information, and comparability of the products.

Processing methods

Product price quote data are pooled by the month, where we exclude products that were only available for a week or less of that month. This helps address issues with data collection, for example, where the collection for a retailer has been affected by changes to the website.

The product size information in our data is not perfect. The web-scraped data do not have product size information at all for many products, and where it is present it is sometimes incorrect, ambiguous, or in nonstandard units. To address this, we have manually searched for size information for several products and applied unit conversion where appropriate. Where we do not have any size information, or the size information is larger or smaller than is plausible, we drop the product from the analysis.

Item-level weights

Items are aggregated according to the Consumer Prices Index (CPI) weights at classification of individual consumption by purpose (COICOP) 5 level. In most cases, each item belonged to a unique COICOP5 category.

For four COICOP5 categories, we have a pair of items belonging in the category. In these cases, the category weight was split evenly between the two items.

Retailer weights

As we use prices from seven high street grocery retailers, we need to use a weight structure to produce weighted average prices across those retailers. We use three main data sources to estimate the market share of each retailer within low-income households: UK grocery market share data, Index of Multiple Deprivation (IMD) Income Domain, and UK supermarket location data.

UK grocery market shares, provided by Kantar, captures aggregate market shares across all income levels.

Therefore, we have developed a method to calibrate those shares to obtain grocery market shares for low-income households.

Income domain of IMD ranks small areas (LSOA) within each country of the UK according to the income deprivation levels. We first match each grocery store to the income decile of the neighbourhood in which it is located. Then, we calculate the proportion of each retailer's stores in the lowest-income decile. Finally, we apply these proportions to the retailer's market share to approximate the market share of each retailer in the lowest income decile areas. These shares are then rescaled to sum to one and used as retailer weights.

Index methodology

To produce the index series at each aggregation level, we use a bilateral Lowe index, which uses quantities in a choice of period to weight each item. The formula for the Lowe Index is given as:

$$P_{FB-Lowe}^{0,t} = \frac{\sum_{i=1}^{N} p_i^t q_i^b}{\sum_{i=1}^{N} p_i^0 q_i^b}$$

where *b* can be any period, or range of periods.

At the first level, to produce the item group index series, we apply this formula for each item group with unit prices for *p* for each product *i* in period *t* and retailer weights for *q* which we assume to be constant over time.

To produce the final index series, at the highest level, we use the item group indices and apply the Lowe formula again, this time treating each item group as a product with index values *p* for each label *i* in period *t* and aggregating according to the CPI weights at COICOP5 for *q*.

9 . Strengths and limitations

Limitations in measuring the lowest price of groceries

The estimates presented here are highly experimental and are subject to great uncertainty. All average estimates will reflect a combination of different price movements for each individual item.

There are several limitations of the analysis. One is that the data are based on prices and product characteristics (including pack size) that have been scraped from retailers' websites. This means that the available products represent the retailer's online catalogue, rather than the range of products available or bought in local stores that month. Moreover, for some retailers, the availability of stock shown online is based on one store and is not representative of the stock across their stores and online more broadly.

We do not have the data to say that a lowest-priced product is actively being purchased by consumers. Although our dataset includes price and product details, we do not have sales or expenditure data; all we know is that it is available to purchase from the retailer's website.

As we wish to focus on the very lowest-cost products, not an average of numerous product prices, the estimates of price change are created from a very small number of price quotes. For each month, figures for each item are based on seven prices at most; these are the products with the lowest unit price from each retailer. This means that the analysis is extremely sensitive to the input data.

With any new experimental process, there may be problems with its implementation. In this case, data have not been collected under ideal conditions. The inability of web scrapers to immediately adapt to changes to retailer websites mean some data were missed on occasion. For example, there is evidence to suggest, over the last year there has been an expanding number of retailers introducing value ranges online, which previously might not have been available online. Even if these ranges were in store the data presented is this article is created using online data web scraped from supermarket's websites only. As data are collected on a daily basis, it is not possible for us to go back and recollect missing data. Where data have been missed, we have developed processes to account for missing prices.

Impact of product substitution on results

Where a product selected in the previous month was not available in the collected data, a substitution was made to the next lowest-priced, similar available item. A substitution was also made where a similar item was available at a lower price. The findings are very sensitive to the approach for substituting products and data on online price quotes and product availability may not reflect instore conditions.

The difference between the cheapest and the next cheapest item is often substantial, and so any type of substitution can have a notable impact on the index and corresponding price change over the year.

The impact of substitution can act to either increase or decrease the price index, at any point in the time series. For example, the lowest-priced or value-range item may not happen to be available at the start of the time period but may come back in stock at the end of the time period. This would act to reduce the index in later periods.

To identify the underlying trend in the lowest prices, without introducing excessive noise from product churn where the range of available products on online stores continually changes - it may be beneficial to constrain the amount of product substitution that is occurring.

One approach to doing this is to limit the allowed substitutions so that they are within a strict percentage price difference range. The benefit of this is that it reflects the fact that spending decisions can change depending on whether a substituted item is substantially more expensive or not. This approach - which is the one we use for the headline results - removes much of the volatility in the index and results in a stronger upward trend movement throughout the year.

We previously used the method that a substitute was not selected if the price was 20% higher or lower than the maximum and minimum prices for the missing item being substituted during the whole time series. However, as the time series extended and with some prices increasing leading to some substitute items not being suitable for the respective period. This created notable volatility in the indices which was not due to price movements but due to missing data.

To reduce this volatility and to ensure items were a suitable substitute, for the respective period, we have updated the substitution method in this release, the 20% threshold for substitution has now been calculated on a rolling basis using the minimum and maximum prices in the current and previous month. Once a product is identified as a substitute for an item it remains eligible for substitution regardless of future price changes. This can cause potentially inappropriate substitution in cases where more expensive products are identified as substitutes while on discount. This updated substitution method means estimates will differ from the previous publication.

This threshold was informed by sensitivity analysis balancing the volatility of the resulting data and the effect on the number of eligible products and the comparability of the substitute item.

A drawback of both these approaches is that it can cause us to miss the entry of cheaper products for an item where the new products come in at a price point far enough below the current product.

Another approach would be to instead allow no constraint when choosing a substitute item; to pick the cheapest product (by unit price within a size band) matching each item, with no regard to the price difference for the substitute item. For some items, this approach could lead to some very expensive products being considered as substitute items, meaning that we would see far greater volatility in the overall time series.

This alternative approach to substitution would result in a notable reduction in the index from February 2022 onwards, reflecting some value items not being available (in online stores) at the very start of the period, with the only available substitutes at that point being much more expensive products. When cheaper value products become available in later months, this alternative method would show reductions in the index relative to the results from our chosen, more constrained, substitution approach.

10 . Future plans

This analysis is part of our current and future analytical work related to the cost of living, which has also included developing our personal inflation calculator to show you how inflation is affecting your household costs.

As we have outlined, since the analysis is based on web-scraped data, there were, inevitably, limitations to the analysis of lowest-cost items that we could carry out.

Our ongoing transformation programme to include new improved data sources and developing our methods and systems for the production of UK consumer price statistics will notably improve our capability to reflect our changing economy and produce more robust, timely and granular inflation statistics for businesses, individuals, and government.

We welcome feedback on this work, which can be addressed to: cpi@ons.gov.uk.

11 . Related links

Consumer price inflation, UK: September 2022
Bulletin | Released 19 October 2022
Price indices, percentage changes, and weights for the different measures of consumer price inflation.

The cost of living, current and upcoming work: September 2022
Article | Released 28 September 2022
A summary of the Office for National Statistics' (ONS') current and future analytical work related to the cost of living.

How is inflation affecting your household costs?
Article | Released 23 March 2022
Use our inflation calculator to see how increases in the cost of living have affected you in the past year.

Transformation of consumer price statistics: April 2022
Article | Released 27 April 2022
Our plans to transform UK consumer price statistics by including new improved data sources and developing our methods and systems for production from 2023.

Household Costs Indices, UK: fourth preliminary estimates, 2005 to 2021
Bulletin | Released 4 May 2022
UK households' experience of changing prices and costs. These are 2005 to 2021 estimates of the Household Costs Indices, a set of measures currently in development.

12 . Cite this statistical article

Office for National Statistics (ONS), released 25 October 2022, ONS website, article, Tracking the price of the lowest-cost grocery items, UK, experimental analysis: April 2021 to September 2022