

# Online weekly price changes methodology

Latest quality and methodology information for online price indices for a selection of food and drink products from several large UK retailers. These are experimental data created as part of the faster indicators release in response to the coronavirus (COVID-19) pandemic.

Contact:  
Rob Kent-Smith  
faster.indicators@ons.gov.uk  
+44 (0) 1633 651 618

Release date:  
4 November 2020

Next release:  
To be announced

## Table of contents

1. [Overview of online price changes data](#)
2. [How we measure online price changes data](#)
3. [Strengths and limitations of online price changes data](#)
4. [Timeline of changes to online price indices data](#)

# 1 . Overview of online price changes data

During the coronavirus (COVID-19) pandemic, we have been providing timely indicators of the effect of the disease on the UK economy and society in our [Coronavirus, the UK economy and society, faster indicators bulletin](#).

The initial faster indicators release on 02 April 2020 included experimental online price indices for several high-demand products (HDPs), covering items such as long-life food, health, and household and hygiene products. These products were chosen at the beginning of lockdown using anecdotal evidence on which products were in high demand from consumers. For more information about the high-demand basket online price changes, please see [Online price changes for high-demand products methodology](#). A full history of price indices for these high-demand products is available in this [dataset](#).

As consumer shopping habits have now started to normalise, this basket has been replaced with a broader selection of food and drink items. Scraping of data from retailers that were previously used in the HDP basket but not in the food and drink basket was discontinued in December 2020.

Experimental price indices for these items are created using daily web-scraped data from several large UK retailers and provide an indication of weekly price change for items in the new basket.

As experimental indices, these data are subject to revisions as we develop our methodology and systems. Any methodological improvements will be detailed in [Section 4: Timeline of changes to online price indices data](#).

We anticipate producing this analysis for the duration of the coronavirus (COVID-19) pandemic to help provide timely data in this period of increased uncertainty. It is not appropriate to compare these online price movements with the headline [Consumer Prices Index including owner occupiers' housing costs \(CPIH\)](#).

## 2 . How we measure online price changes data

Prices are scraped daily from several large online UK retailers (typically supermarkets and other prominent high-street chains with an online presence). Prices have been collected from 01 June 2020 for selected items chosen to form the new basket. A full list of items is available in the [dataset](#).

The items in this new basket have been chosen to align with items collected in our official consumer price statistics in the food and non-alcoholic beverages, and alcohol categories. Items in our official consumer price statistics are chosen to be representative of the market and consumer purchasing behaviour, therefore this basket follows the same weighting principles. For more information on this process, please see our [Consumer Prices Technical Manual](#).

However, due to the increase in product coverage, the definition of items collected in the weekly online price changes release are typically broader in specification than those collected in our official consumer price statistics. This allows for more accurate semi-automated classification of the products once scraped.

An average weekly price is calculated for each unique product, and the change in the average weekly price over time is used to construct the price index.

## Method

Because of the nature of these data and the impact that daily changes in product availability can have on chain drift within a chained Jevons series, we have chosen to apply a movement splice GEKS-Jevons index to these data. The movement splice GEKS-Jevons method is essentially equivalent to the GEKS method but it uses a rolling window of five weeks to ensure the data remain relevant to the current time period.

For more information on the movement splice GEKS method and our wider research into suitable index methodology for alternative data sources such as web-scraped data, please see our latest project update [new index number methods in consumer price statistics](#).

## Higher-level indices

The faster indicators publication will focus on category and all-items level indices, although the item indices will also be published in a dataset.

Our category indices provided in the release are each derived from a number of lower-level items. These item indices are aggregated together using expenditure weights that represent how much consumers typically spend on them. Given a lack of timely expenditure information at this item level, we have used the Consumer Prices Index including owner occupiers' housing costs (CPIH) item level weights to inform our estimates. These weights are derived from a time before our base period; therefore we are using a Lowe formula for calculation of our higher-level indices. Our all-items index is based on a similar aggregation of our category level indices.

To understand the main drivers towards the change in the weekly rate at the category level, we look at each items' contribution to this change. Contributions are calculated using a weekly version of the formula used to calculate contributions to the month on month change for our official estimates of consumer price statistics. Please see Section 10.7 of our [Consumer Prices Technical Manual](#) for more information.

## 3 . Strengths and limitations of online price changes data

### Comparability with other sources

It is not appropriate to compare the online price movements detailed in the [Coronavirus, the UK economy and society, faster indicators release](#) with the headline [Consumer Prices Index including owner occupiers' housing costs \(CPIH\)](#).

The CPIH is produced using different methods, data and quality thresholds, and incorporates a broader range of goods and services, such as housing. The online price movements data are experimental and are produced using experimental web-scraping techniques.

### Strengths of the experimental online price indices

The data present an indicator of the price movements of a basket of products during the coronavirus (COVID-19) pandemic.

The data are available on a weekly basis, to allow week to week comparisons of price movements in these products.

The data are available four days after the reference period. This makes them a timely indicator for users.

## Limitations of the experimental online price indices

No comparable replacement or quality adjustments have been applied, which may lead to downward bias in the indices.

Because of the nature of web-scraped data, there are no expenditure weights available at the product level. Therefore, there is no accounting for consumer substitution within these items. As we follow the same unique product over time, any substitution towards cheaper or more expensive products will not have an impact on the index, unless these types of products follow a different price change trend.

For retailers that have a clear label indicating that a product is out of stock, we have removed those products from the analysis. However, other retailers may not have a clear indication of this, so there may be products currently out of stock that have still been included. If the prices of these items do not change, this could cause the index to remain static.

Limited sample size and daily changes in product availability may result in volatility at the item level. In particular, the size of the sample means that sometimes single retailers can contribute to substantial movements at the item level.

The number of products and retailers we are collecting mean that it is not possible to individually scrutinise every price quote, and we use semi-automated methods to classify products and remove outliers. In our quality assurance procedures, we may identify products that have been misclassified. When this occurs, we will update the time series to remove or include these products.

## 4 . Timeline of changes to online price indices data

The publication of online price indices data for high-demand products was included in the [first faster indicators publication on 2 April](#). The publication of the new basket was first included in the [faster indicators' publication on 10 September](#).

As experimental indices, these data are subject to revisions as we develop our methodology and systems. Due to the methods used to classify products and our regular quality assurance procedures, we may make minor improvements to the classification of products over time, which could have a small impact on the historical index series. These minor changes will not be flagged in this section. We recommend that users use the most recent dataset when looking at historical data.

We began collecting data from a new retailer on 12 October 2020. The historical time series has been updated accordingly. The impact of these changes is small at the aggregate basket and category level, although there were some larger changes for individual items, such as battered fish.

This section will be updated over time to provide further detail on any significant revisions.

The updated historical time series for all individual items are published in a [dataset](#), updated weekly.