

Article

Transformation of consumer price statistics: November 2021

Our plans to transform UK consumer price statistics by including new improved data sources and developing our methods and systems for production from 2023.

Contact:
Sofia Poni
cpi@ons.gov.uk
+44 1633 456900

Release date:
9 November 2021

Next release:
To be announced

Table of contents

1. [Main changes](#)
2. [Overview](#)
3. [Alternative data sources](#)
4. [Why we want to use alternative data sources](#)
5. [Current methods of data collection](#)
6. [Changes in how we process data](#)
7. [Implementation plan](#)
8. [Future developments](#)
9. [Related links](#)

1 . Main changes

- We are undertaking a programme of transformation across many areas of our price statistics, including identifying new data sources, improving methods and developing systems.
- We currently collect the majority of prices from physical stores; however, in the future we envisage our consumer price statistics will be a mix of scanner data taken directly from checkouts, information scraped from shops' websites, administrative data, and traditionally collected data for smaller or independent businesses.
- The developments outlined are part of a continuous programme of improvement for consumer price statistics; our ambition is to bring in new data sources for further areas of the inflation basket and continue to improve these statistics over the coming years.

2 . Overview

Accurate measures of inflation play a vital role in business, government and everyday life. From rail fares to taxes and pensions, financial transactions and many other areas of our lives are regularly adjusted to reflect the change in prices over time. It is crucial to measure these changes in price as accurately as possible.

We are currently undergoing transformation across many areas of our price statistics, including identifying new data sources, improving methods and developing our statistical processing systems.

As part of this transformation, we are updating the way we collect and process price information to reflect our changing economy and produce more robust, timely and granular consumer price statistics for businesses, individuals and government.

3 . Alternative data sources

We are introducing several new data sources to help us transform our collection.

Scanner and transaction data

Scanner data are collected by retailers at the point of sale. Scanner data provide us with significantly more information on the number and type of products sold, allowing us to more accurately reflect changing consumer spending patterns.

Our ambition is to collect billions of prices every month from the UK's leading retailers, which will improve our understanding of how prices of products, as well as the number of each product sold, are changing in the UK economy.

We are engaging directly with the UK's largest retailers to gain access to scanner data. Importantly, the data will show total weekly figures regarding what retailers have sold, not what individual people have bought.

Many of these retailers are already sending us regular feeds of scanner data that together account for almost 50% of the food and drink market. We are continuing the process to acquire data from further retailers to continually increase our market share coverage.

We also have transaction level data for rail fares in Great Britain. These contain approximately 2 million transactions per day, which give us a much better picture of how rail fares are changing. They include total ticket cost and price-defining attributes, such as origin and destination station, ticket type and class.

Web-scraped and web-provided data

Web-scraped price data are collected from retailers' websites and can provide a wealth of additional product information about online prices, such as product descriptions. For example, as well as obtaining the price of a laptop, we can collect information such as RAM and processor speed, which help us understand how the quality of products changes over time, an important factor when calculating inflation.

We receive web-scraped data covering areas such as clothing, electronic items and package holidays. There are no historical series available with these data (unlike scanner data) so we have been building up a sufficient data time series before they can be used in official measures.

Working with a third-party data provider we now also have access to used car prices data provided by an online marketplace. This gives us a data time series of price history and a significant increase in model coverage. Used cars are an area of the basket where we have witnessed significant shifts in consumer trends, that highlights the ever-growing importance of using alternative data sources to enhance our coverage of the used car market.

Administrative data

Administrative data are information collected by government departments and other organisations (public, private and third sector) primarily for operational reasons rather than statistical ones. They can provide greater coverage than survey data alone.

We will continue to [use administrative data sources](#) that are already currently used for some areas of the consumer basket, such as for owner occupiers' housing costs and private rents. In addition to these existing data sources, we are also now receiving [rental information at a microdata level](#) from the Valuation Office Agency, Welsh Government, Scottish Government and the Northern Ireland Housing Executive.

4 . Why we want to use alternative data sources

Alternative data sources provide many benefits compared with our current sources, including improved product coverage, high frequency of collection, as well as potential cost savings. Scanner data also provide additional information such as expenditure per product, while web-scraped data contain a rich source of product information that is useful for things like accurate classification and determining quality. Scanner data also have the potential to provide greater regional coverage of prices and expenditure such as, for example, regional inflation measures.

5 . Current methods of data collection

We will continue to use traditionally collected price data and our existing administrative data in Consumer Prices Index (CPI) and Consumer Prices Index including owner occupiers' housing costs (CPIH) when they cannot be replaced by scanner or web-scraped data, such as for small independent shops who do not have a website.

We envisage that in future our consumer price statistics will be a mix of scanner, web-scraped, administrative and traditionally collected data.

6 . Changes in how we process data

These new data sources will result in hundreds of millions more price quotes being processed each month. To maximise their benefit, we will also be making changes to our methodologies and systems.

For the Consumer Prices Index including owner occupiers' housing costs (CPIH) and Consumer Prices Index (CPI), changes are required at the lowest level of aggregation to integrate these new data, while ensuring that they are appropriately represented within our price indices. We are also enhancing some of our current methods for traditionally collected data. This is discussed further in the latest article on [Introducing alternative data into consumer price statistics: aggregation and weights](#).

For any changes in methodologies and systems we will undertake extensive stakeholder engagement and continue to publish [biannual methodological articles](#) and impact analyses.

7 . Implementation plan

We plan to include data from alternative sources in our headline consumer price statistics from Quarter 1 (January to March) 2023, dependent on continuous engagement with users, as outlined in [Section 8, Timeline](#).

The first divisions we aim to include alternative data sources for are food and non-alcoholic beverages (classification of individual consumption by purpose: COICOP 01), alcohol and tobacco (COICOP 02) and transport (namely rail fares) (COICOP 07.3). These divisions will be partially based on scanner and transaction data, in conjunction with traditionally collected data.

Throughout 2020 and 2021, we have been working on further developments of the systems for web-scraped and scanner data to enable research and impact analysis, alongside research into the methods needed to produce high-quality indices using web-scraped and scanner data. More detail can be found in our [Consumer Prices development work plan](#).

The next phase of the project will run during 2022 and involves the regular publication of aggregate experimental indices including scanner data (groceries and rail fares) in conjunction with traditionally collected data, as well as a range of impact analyses.

As well as presenting the aggregated experimental indices, the impact analyses will decompose the effects on the contributions to the 12-month inflation rate at divisional level. Further analysis of these areas of the basket (of consumer goods and services) will also be included, which will be published as a quarterly release.

8 . Future developments

The changes outlined are the initial stages of a continuous programme of improvement for consumer price statistics. Our ambition is to bring in new data sources for further categories and continue to improve consumer price statistics over the coming years.

Throughout each phase, we will also be liaising regularly with our [Advisory Panels for Consumer Price Statistics](#), our users, and the Office for Statistics Regulation, to ensure that our plans for consumer price inflation measurement are appropriate for improving the quality of our statistics and meeting our ongoing user requirements.

Timeline

The following timelines are subject to continued systems development, research and impact analysis to ensure the quality of our statistics, which is our priority. Decisions will be made through continuous engagement with our stakeholders:

- 2020: Continued research into the methods required to process alternative data sources
- 2020 to 2021: Development of systems for processing alternative and traditional data sources
- 2021: Application of methods and impact analyses for priority items
- 2022: Recommendations of methods for each item and data source and stakeholder engagement
- 2022: Publication of impact analyses and experimental estimates of Consumer Prices Index (CPI) and Consumer Prices Index including owner occupiers' housing costs (CPIH) incorporating alternative data sources for groceries (classification of individual consumption by purpose: COICOP 01 and 02) and rail fares (COICOP 07.3)
- 2022: Continued research and application of methods, systems development and impact analyses for additional priority items
- Quarter 1 (January to March) 2023: Alternative data sources for groceries (COICOP 01 and 02) and rail fares (COICOP 07.3) used in aggregate measures of consumer price statistics
- Quarter 1 (January to March) 2024: Alternative data sources for additional priority categories, such as web-scraped clothing data, used in aggregate measures of consumer price statistics
- 2025 and beyond: Rolling programme of improvements to use of alternative data sources, including use of additional data in a category, roll-out of alternative data sources for new item categories, and methodological and systems improvements required for the use of alternative data sources

9 . Related links

[Introducing alternative data sources into consumer price statistics: May 2020](#)

Article | Released 27 May 2020

Our plans to include alternative data sources into the production of consumer price statistics.

[Consumer prices development plan: updated March 2021](#)

Article | Released 16 March 2021

The proposed development plan for consumer price statistics considers a range of priorities, including the rationale for these priorities and their potential impact.

[Research and developments in the transformation of UK consumer price statistics: November 2021](#)

Article | Released 9 November 2021

The third in a series of biannual articles to update users on our research to modernise the measurement of consumer price inflation in the UK.

[Consumer price inflation, UK:](#)

Bulletin | Monthly

Price indices, percentage changes, and weights for the different measures of consumer price inflation.