

Article

Consumer prices development plan: updated March 2020

The proposed development plan for consumer price statistics considers a range of priorities, including the rationale for these priorities and their potential impact. Resource requirements and a broad timeline are included to provide users of prices statistics with a greater understanding of the planned development work.

Contact:
Chris Payne
CPI@ons.gov.uk
+44 (0) 1633 456 900

Release date:
19 March 2020

Next release:
To be announced

Table of contents

1. [Introduction](#)
2. [Organisational context](#)
3. [Priorities](#)
4. [Work programme for alternative data sources](#)
5. [Work programme for the range of consumer price statistics](#)

1 . Introduction

The Office for National Statistics (ONS) has a programme of research and development aimed at improving and maintaining its range of consumer price inflation statistics – ensuring that they continue to meet user needs, make use of new and innovative methods and data sources, and follow international best practice.

The article, [Measuring changing prices and costs for consumers and households, proposed updates: March 2020](#), sets out more detail on the “use cases” for each of our main inflation measures – the Consumer Prices Index including owner occupiers’ housing costs (CPIH), the Consumer Prices Index (CPI), the Household Costs Indices (HCIs), and the Retail Prices Index (RPI).

2 . Organisational context

The strategy for UK statistics, [Better statistics, better decisions](#), which runs to April 2020, sets the collective mission for the official statistics system as, “high quality statistics, analysis and advice to help Britain make better decisions”. This strategy is reflected in the [ONS Economic Statistics and Analysis Strategy: financial year ending 2019](#), of which consumer price statistics are a part. The development plan presented in this article therefore reflects how consumer prices development work contributes to the strategy for Economic Statistics and the wider UK statistical system, through quality and relevant price statistics that are produced efficiently to meet customers’ needs, keeping pace with evolving methods, sources and digital processes.

Items on the development plan are prioritised through discussion at our [Advisory Panels on Consumer Prices](#) every May,¹ and are assigned a high, medium or low priority. High priority items will be resourced as a priority, and we expect to make good progress in these areas. High priority work streams would only be stopped if it were necessary to ensure the publication of our main outputs. Medium priority items will be pursued as resource allows; however, some medium priority work may be slowed down or stopped altogether to allow progress on high priority items to be made. Low priority items will only be taken forward where resource is available to do so, and this will not prevent high or medium priority workstreams from progressing.

Notes for: Organisational context

1. September in future years

3 . Priorities

The work programme for consumer price statistics was first included in the 2015 consultation on consumer price statistics. The current work programme has since been updated to reflect the outcome of the review, and also to reflect decisions reached by the National Statistician following the consultation. It has also been updated to reflect the views of the Advisory Panels on Consumer Prices following discussion by both the Technical and Stakeholder Panels every May between 2016 and 2019. The programme and priorities also reflect the priority placed on the use of alternative data sources by Professor Sir Charles Bean’s [review of economic statistics](#) (commonly referred to as the Bean Review).

We expect that our top priorities for the next few years will be to continue to develop the range of consumer price measures to meet the three user needs identified by the National Statistician – a set of measures consistent with recognised economic principles, measuring the household experience of changing prices and costs, and consulting on the future of the Retail Prices Index (RPI). Our priorities, therefore, are to:

- embed the CPIH as the lead inflation measure in the UK by providing users with a longer time series and more granular data, continuing to provide information and assurances on sources used to compile the Consumer Prices Index including owner occupiers' Housing costs (CPIH), and improving user understanding of the measure
- continue to produce and improve the Consumer Prices Index (CPI) so that we have a measure that is comparable internationally
- work to develop a set of indices that reflect inflation as experienced by household groups
- consult on the future of the Retail Prices Index (RPI)
- improve all of our consumer price statistics through the use of alternative sources of price and transaction data

4 . Work programme for alternative data sources

We are currently working through a comprehensive transformation programme for consumer price statistics to modernise their measurement and make better use of data and methods that are becoming increasingly available to us.

At a high level, this involves obtaining robust sources of alternative data, development of statistical systems to work with these data, and methodological research to effectively classify, validate and construct high quality price indices from new data sources. These new data sources will be used in conjunction with traditionally collected data to improve the accuracy, efficacy and representativity of consumer price inflation statistics.

The data sources that we are investigating are web-scraped data (automated data collection from retailer websites) and scanner data (point-of-sale expenditure and quantity data provided directly by retailers). More information can be found regarding these data sources in our article [Introducing alternative data sources into consumer price statistics](#).

This transformation will be the largest change to consumer price statistics in a generation, and the scale and importance of this work should not be underestimated. We will be reliant on developments in many areas, including the use of new technology platforms and the willingness of retailers to provide us point-of-sale data. Because of the scale of this work, the project has been separated into its own work programme, to allow for more detail to be provided about the various work streams.

Our ambitious roadmap sets out our plans to incorporate alternative data sources into our headline measures of consumer price statistics by 2023. We plan to do this using a three-phased approach:

- Phase 1 (Research: 2020) involves developing systems and methods for use with alternative data sources, alongside traditional sources and methods.
- Phase 2 (Application: 2021) involves us applying these methods to specific categories as [prioritised with our Stakeholder Advisory Panel on Consumer Prices \(PDF, 92.66KB\)](#).
- Phase 3 (Engagement: 2022) involves the release of quarterly experimental estimates of the impact of alternative data sources on consumer price statistics, as well as engagement with stakeholders and users about these changes.

The impact of the new data sources and methods will first be available in the published figures in the February 2023 UK consumer price statistics, published in March 2023.

4.1 High priority

High priority items are the cornerstone of the development programme and, if necessary, will be prioritised over the delivery of medium and low priority items.

The workstreams listed summarise the research into new methods and systems that is required before we can process alternative data sources for the purposes of producing consumer price statistics in Quarter 1 (Jan to Mar) 2023. Where relevant, intermediate implementation dates are also included.

Developing a processing pipeline

Overview

The pipeline to produce consumer prices indices using web-scraped data is being built on our in-house distributed system – the Data Access Platform (DAP). The platform is based on a very powerful cluster of computers and provides users with many software tools to store and analyse data.

The pipeline is being designed using a suitable flexible framework that can be applied to all items and data sources in the future consumer basket.

Implementation

- Functionality of pipeline expanded to include locally collected data - Quarter 3 (July to Sept) 2020
- Complete development of our IT system - Quarter 4 (Oct to Dec) 2021
- Use processing pipeline for parallel run alongside existing production system - Throughout 2022

Framework for assessing the quality of consumer prices indices produced using alternative data sources

Overview

This work will summarise the properties of a desirable index number method and provide recommendations on how a final index number method could be selected for different prioritised item categories, incorporating alternative data sources.

The recommendations from this will feed into our final decision on which method/s to choose for implementation.

Implementation

- Final recommendations due on most appropriate index number methods to use given different market and pricing behaviours - Quarter 2 (Apr to June) 2020
- Shortlisted index number methods built into processing pipeline for further testing - Quarter 4 2020
- Final recommendation on index number methods for each prioritised item category - Quarter 4 2021

Classification techniques

Overview

The classification project looks at automatically classifying products to a specific item category. This work will recommend which methods are suitable for our prioritised item categories and for different data sources.

This work will also include a review of existing item definitions, as some look to be too narrow for automatic classification techniques to work to a suitable level of accuracy (for example, the item definition for women's blouses specifies whether the blouse should open fully, information that is not readily available in item descriptions scraped from retailer websites).

Implementation

- Case studies of different classification methods that are suitable for different types of product category and different data sources - Quarter 4 2020
- Classification pipeline built - Quarter 4 2020
- Final recommendation on classification methods for each prioritised item category - Quarter 4 2021

Expenditure weights for web-scraped data

Overview

One of the limitations of web-scraped data is that it does not provide information on expenditure or quantities of product bought. This work will identify if the lack of expenditure weights at the product level introduces any bias into any index based on web-scraped data, and if we can approximate expenditure weights using alternative data sources like page rankings.

Implementation

- Final recommendations due on if/how expenditure weights can be found for web-scraped data at the product level - Quarter 4 2020

Product grouping

Overview

In the current methodology, the price for an individual product is followed over time and compared back to the price of the same product in the base period. An alternative approach would be to follow the average price of a defined group of homogeneous products instead. Research has shown this to be a viable alternative for categories such as clothing, which experience high rates of product churn over time. This will feed into how we define a unique product in the pipeline for identified categories.

Implementation

- Recommendations on different product grouping methods and what scenarios they should be used in - Quarter 4 2020
- Product grouping pipeline built - Quarter 4 2020
- Final recommendations due on whether product grouping methods should be used for each prioritised item category - Quarter 4 2021

Expenditure weights for different data sources and retailers

Overview

This work will recommend methods and suitable data sources that will allow us to aggregate together alternative data sources with traditionally collected data. For example, traditionally collected data for bread from local bakeries alongside scanner data for loaves bought from large retailers. This will feed into the aggregation part of the processing pipeline.

This work will also be an extension of the improvements to shop-type weights introduced in Quarter 1 2020. This work will incorporate stratification using data from the Annual Business Survey (ABS) for different types of shops.

Implementation

- Update shop type weights in the CPI and CPIH using ABS data - Quarter 1 2020
- Final recommendations on expenditure weights for aggregation - Quarter 4 2021

The impact of product returns, discounts and product relaunches on alternative data sources

Overview

The issue of returns affecting expenditure weights for particular categories may affect how we can use expenditure weights in a final item index.

Take-up rates from discounts observed in scanner datasets may be applied to web-scraped and traditionally collected data to ensure consistent treatment of discounts.

Product relaunches need to be identified to ensure that quality changes in products are captured and appropriately adjusted for over time.

Implementation

- Final recommendations on treatment of returns, discounts and product relaunches - Quarter 4 2020

Using data elsewhere in basket pre-2023

Overview

There are areas in the inflation basket that may benefit from the use of alternative data sources prior to our main implementation milestone in 2023.

The hedonic modelling process for technological goods currently involves the use of hundreds of manually collected data to produce regression models. Web scraping therefore has the potential to improve the efficacy of data collection for use in hedonic modelling, without feeding directly into the indices produced.

There are also some categories in the inflation basket where it may be appropriate to replace existing data collection procedures with web-scraping in-house. For example, we are developing a “robot tool” that sends an email to price collectors when a website changes. This may be useful in cases where we know prices to be relatively static, such as passport fees.

We are also building more advanced web scrapers to scale up our capability to web scrape retailers’ websites.

Implementation

- Supplement hedonic collection “test data” with data from alternative data sources -Quarter 4 2020
- Parallel run advanced web scrapers for motorbikes with live collection - Quarter 4 2020
- Build and parallel testing of robot tool for items with relatively static pricing structures within the basket - Quarter 4 2021

Expanding pipeline functionality

Overview

Improvements need to be made to some of the existing functionalities within the current pipeline. For example, the current functionality uses very basic outlier detection methods (for example, minimum and maximum checks) and imputation approach (price carry-forward method).

Better methods need to be understood and developed, in line with international best practice in price indices and big data.

Implementation

- Recommendation of outlier detection methods for use with alternative data sources - Quarter 4 2020
- Recommendation of imputation methods for use with alternative data sources - Quarter 4 2020
- Review of pipeline functionalities and any further research and/or improvements that are needed - Quarter 4 2020

5 . Work programme for the range of consumer price statistics

5.1 High priority items

High priority items are the cornerstone of the development programme and, if necessary, will be prioritised over the delivery of medium and low priority items.

Developing Household Costs Indices (HCIs)

Overview

The concept of a Household Costs Index (HCI) was first proposed by Astin and Leyland (as the Household Inflation Index, HII), culminating in a paper¹ submitted as a response to the 2015 consultation on consumer price statistics. Suggested differences from existing measures of price change include the potential inclusion of asset prices and interest payments, plus giving each household's expenditure equal weight. Following the consultation, the National Statistician decided that the HCIs could serve as an important complement to the suite of consumer prices indices.

Work has now commenced to develop these indices, releasing focused analytical articles throughout 2017 and beyond that will help shape the production of the final measure. We have subsequently engaged with our advisory panels on a number of development issues. Initial experimental indices were [published](#)² in December 2017 with an update in [April 2019](#), but there are a number of concepts that will need to be explored further. In his [statement](#) of 28 June 2019 the National Statistician confirmed plans to produce further experimental publications on an annual basis throughout the development phase, moving to a test run of quarterly production in 2022. Moreover, once National Statistic status for HCIs has been achieved, we will look to develop a HCIs variant (the HCIs – Capital, or HCICs) which will additionally include capital payments for owner occupied housing.

Following the 2015 consultation on consumer price statistics, the National Statistician reached the decision that we should produce comparable measures of income and price change for different household groups on an annual basis in one publication. We will engage with users to shape further development of the measures.

Much consideration has gone into how comparable measures of income and price change should be produced, and work in the area suggests there may be a case for a “family of indices” to deflate income using different definitions. As such Household Costs Indices may be appropriate to be matched to a microeconomic measure of income. This item ties in with [ESCoE's work stream](#) on democratic measures of income growth which aims to make a wider range of measures available to policymakers.

Implementation

- Release third revision of the experimental Household Costs Indices - Quarter 2 (Apr to June) 2020
- Supporting methodological article for the third revision of the experimental Household Costs Indices - Quarter 2 2020
- Begin test run of quarterly HCIs production - 2022

Developing a historical series for the CPIH

Overview

The Consumer Prices Index including owner occupiers' housing costs (CPIH), the lead measure of inflation, is the most comprehensive measure of consumer prices inflation, given its inclusion of owner occupiers' housing costs (OOH). The current time series extends back to 2005, which is when the rental data sources used to calculate the OOH component for England begin.

To increase the usability of the CPIH and provide a longer commentary on the impact of owner occupiers' housing costs on inflation, a historical series will be modelled for the CPIH.

In January 2018 both Advisory Panels on Consumer Prices considered a proposed methodology for the historical series, and the series from 1988 to 2004 was published in December 2018. The series from 1947 to 1987 is being investigated further.

Implementation

- Further methodological development work on the early part of the series - Quarter 3 (July to Sept) 2020
- Publish the CPIH historical series for 1947 to 1987 - Quarter 4 (Oct to Dec) 2020

5.2 Medium priority items

Medium priority items form an important part of the work programme but if necessary, delivery may be delayed for high priority items.

Developing measures of accuracy for the CPIH

Overview

Because of the complex survey design, calculating standard errors for the CPIH and specifically the growth in the CPIH is very challenging. There are two dimensions to the sampling – selection of items and selection of outlets. The sampling involves purposive sampling of both items and outlets. The CPIH weights come from a variety of sources. Some are administrative sources, some are survey based. They are then put through national accounts balancing. This all makes the estimation of sampling errors difficult.

Also, CPIH weights are price updated using movements in the appropriate CPIH index, so there is an interaction between weights and prices. Annual rates of change in the price index are ratio estimates, usually over a chain link, so there may be correlations between prices in successive periods that need to be taken into account.

An article on the effect of variance in the weights of the CPIH was published in autumn 2017,³ and work on the effect of variance in the prices was reviewed by the Advisory Panels on Consumer Prices in 2018. The work will be prepared for submission to an academic journal in 2020.

Implementation

- Submission of work to estimate the variance in CPIH - Quarter 2 2020

Review of quality adjustment and monitoring of quality change

Overview

One of the more difficult issues in producing consumer price inflation statistics is the accurate measurement and treatment of quality change because of changing product specifications. The Johnson Review considered quality change and recommended we provide more information to explain how quality change is monitored. In response to the 2015 consultation on consumer price statistics, some users suggested that we should devote more resources to improving and monitoring methods of quality adjustment. In particular, the quality adjustment of services could be given more consideration. The Bean Review⁴ also considered quality change and found that the issue is likely to grow in importance with the “spreading tentacles of the digital revolution”.

To address these findings, a project has been initiated to review quality adjustment methods used in consumer price inflation statistics and provide more detail on how quality adjustment is monitored.

This item ties in with [ESCoE's work stream](#) on measuring activity in services sectors, as the project aims to investigate the deficiencies in the current measures of services activities for the UK and how might they be improved.

Implementation

- Monitoring tool for between-year quality changes - Quarter 2 2020
- Implementation of live quality adjustment monitoring - Quarter 1 (Jan to Mar) 2021
- Analysis on forced quality adjustment and practical applications of the findings - Quarter 1 2022

Regional indices

Overview

Amongst other user interest, in February 2018 an Economy, Jobs and Fair Work Committee of the Scottish Parliament recommended⁵ that a price index for Scotland should be given priority. Regional price indices could also improve the granularity of other high priority work such as the Household Costs Indices and subgroups on a CPIH-consistent basis.

The regular collection of prices for consumer price inflation statistics is optimised for measuring inflation at the UK level. Prices are collected locally in 141 locations spread across Scotland, Wales, Northern Ireland and the nine regions of England. As a result, the number of locations visited per region is small, making the data less suitable for regional indices. Additionally, the prices for many products are collected centrally with no regional breakdown. The viability of using the existing price data to produce more geographically disaggregated price indices was assessed in a feasibility report published in November 2017, and the use of small area estimation to improve regional expenditure estimates was explored in a further paper in February 2019.

The University of Strathclyde will also deliver a report detailing the research done in 2019. The report will (among other things) focus on the price data used in the regional CPIH. It will “assess the temporal stability of the regional price quotes and the capacity for price estimation to be improved using model-based methods”, along with providing recommendations and proposals for future work and possibilities for regular publication.

Implementation

- Ongoing development work towards the production of regional price indices - Quarter 2 2020

Further quality assurance of the CPIH

Overview

As part of the work to seek re-accreditation for the CPIH as a National Statistic, we developed a Quality Assurance of Administrative Data (QAAD) document for our consumer price statistics.⁶ The QAAD highlighted a number of areas where the quality assurance for some data sources could be improved.

We continue to seek the required assurance for these sources and will aim to update the QAAD biennially.

Implementation

- Update to the QAAD for consumer price statistics - Quarter 3 2020

Improvements to CPIH methodology

Overview

There are a number of ways in which the owner occupiers' housing costs (OOH) component of the CPIH could be further improved. We intend to explore these potential developments to further improve the effectiveness of the CPIH as a measure of inflation. These are described in this section.

When the OOH component in the CPIH was developed in 2011 using administrative data sources for England, Wales and Scotland, comparable rental data for Northern Ireland were not suitable. At the time, the Northern Ireland Housing Executive (NIHE) received private rental data biannually, covering the Belfast Metropolitan Region only. As a result, the existing CPI private rental data series for Northern Ireland has continued to be used. Since then the coverage across Northern Ireland has improved and the data are now available monthly. We have received an extract of these data from NIHE and started analysis with the ultimate aim of producing a rental index suitable for inclusion in OOH.

Currently, dwelling stock data from the Ministry for Housing, Communities and Local Government (MHCLG) are used to mix-adjust rental data to reflect the OOH market. However, the property type split is not available on a regional basis, so the same property type split is applied across all regions. We will explore other potential sources of data to improve the stratification.

The Johnson Review⁷ identified that recent research into the measurement of rental equivalence has suggested using a flow measure (new lets only), rather than a stock measure (new and existing lets). We will also explore the appropriateness of using this new approach to measuring rental equivalence costs.

Implementation

- Report on the findings of implementing the new rental data source for Northern Ireland in OOH - Quarter 4 2020
- Present findings from investigation into stock and flow measures of rental equivalence - 2020
- Improved property type split for stratum weights - 2021

Improvements to elementary aggregate indices

Overview

In 2010, we made a number of changes to the methodology used to collect clothing prices. These changes meant that the gap between the RPI and the CPI, which use different formulae at the lowest level of aggregation, ⁸ widened.

The work will consider recommendations from the Johnson Review to review and publish the criteria for formula selection at the lowest level of aggregation.

This work links to the high priority workstream from Section 4.1 on a framework for assessing the quality of consumer prices indices produced using alternative data sources.

Implementation

- Review of criteria for applying elementary aggregate formula, and proposals - Quarter 1 2021
- Impact assessment of proposed index criteria - Quarter 3 2021

5.3 Low priority items

The delivery of low priority items may be delayed or even stopped to ensure the delivery of high and medium priority items.

Review the existing methodology for reconciliation between the CPIH and RPI

Overview

With the move towards making the CPIH the preferred measure of inflation, we have developed a reconciliation between the CPIH and RPI, on a consistent basis with that currently published for the CPI and RPI. In the longer term, alternative approaches will be investigated, using detailed item-level information.

Implementation

- Experimental improved reconciliation method - Quarter 2 2022

Improvements to OOH(NA)

Overview

We currently produce an experimental net acquisitions index for Eurostat, as part of a pilot to incorporate owner occupiers' housing (OOH) costs into the HICP. Currently there is no weight for the component "existing dwellings new to the household" sector, which means that it is given a zero weight in the aggregation. Moreover, the weight for the "acquisition of new dwellings" component includes new dwellings outside the OOH sector (that is, the weight is gross acquisitions, rather than net). We will explore data sources that could be used to improve this experimental index.

Implementation

- Improved methodology for the net acquisitions index - 2022

Inclusion of FISIM in the CPIH

Overview

Financial intermediation services indirectly measured (FISIM) are included in the national accounts measure of household final consumption expenditure (HHFCE) but are not currently included in consumer price indices. The scope of the CPI, which is governed by European legislation, is drawn from the same source as HHFCE and the exclusion of FISIM is one of the biggest differences between the two. As the CPIH is not bound by the same legislation, we will consider the suitability of including FISIM as a proxy for the service charge that households pay to banks. A review will include consideration of conceptual appropriateness, international practice and methodology used to calculate a price index for FISIM.

Implementation

- Feasibility study on the inclusion of FISIM in the CPIH - 2022

The extent to which consumers substitute between outlets

Overview

The Johnson Review recommended that we should research the extent to which consumers substitute between outlets. For example, this would capture how prices for the same goods have changed with the move from corner shops to supermarkets and from supermarkets to online providers.

Implementation

- Report on the extent to which consumers substitute between outlets - 2022

Notes for: Work programme for the range of consumer price statistics

1. [Towards a Household Inflation Index](#)
2. [Household Costs Indices, UK: preliminary estimates 2005 to 2017](#)
3. [Survey Methodology Bulletin, autumn 2017](#)
4. Details of the [Bean review and the final report](#).
5. [How to make data count: improving the quality and coverage of economic statistics](#)
6. [Quality assurance of administrative data used in consumer price inflation statistics](#)
7. Details of the [Johnson review and the final report](#)
8. [CPI and RPI: increased impact of the formula effect in 2010](#)