ADVISORY PANEL ON CONSUMER PRICES – STAKEHOLDER

Scope of groceries scanner data parallel run and impact analyses

Purpose

1. This paper provides Panel members with an overview of the ONS's proposed quality assurance plans for incorporating grocery scanner data into our headline measures of consumer price inflation in March 2025
2. In particular, it sets out our plans in relation to two key quality assurance components: the impact analyses to be produced and the scope of the parallel run to be undertaken in the months leading up to March 2025

Actions

3. We ask for the panel's view on whether:
    a. The quality assurance plan as set out here creates sufficient confidence to allow the ONS to go live with grocery scanner data in March 2025.
    b. There is a need for any other internal or external quality assurance mechanisms and/or an extended parallel run period.

Background

4. Groceries scanner data will be integrated with traditional data sources in the production of UK consumer price statistics from March 2025, using the approach set out in our past publications, as summarised here.

5. Scanner data will never completely replace the traditional, manual collection as there will remain a need for price collection for retailers and services who do not have a means of providing data or an online presence that allows web-scraping. Combining data from these different sources will require significant methodological changes to, for example, our approach to aggregation and imputation. To enable these methodological changes and to move away from using legacy systems, the processing of locally traditionally collected data (i.e. the manual collection in shops) and elementary aggregation across scanner and traditional data is being moved to a new platform.

6. To ensure the quality of the resulting new indices and the stability and robustness of the data, systems and processes that underpin them, we are implementing an extensive quality assurance plan over the next year. This paper summarises this plan, with a particular focus on the proposed parallel run and the impact analyses to be produced and disseminated.

Quality assurance overview

7. The parallel run, wider testing activities and the impact analyses are often conflated and it can be confusing as to how these inter-relate. Table 1 below sets out the different quality assurance tools and processes we have put in place to provide confidence that all user acceptance criteria can be satisfied. Given how much work has already been done to develop and test methods, systems, pipelines and applications, many of these acceptance criteria have already been partially met, but the impact analyses and parallel run activities will be at the core of our quality assurance in 2024/25, as set out in detail below.

Table 1. Quality assurance matrix

	Quality assurance activity									
Acceptance Criteria	**On-going data checks and monitoring**	**Internal & external review, advice and approvals (incl. APCP)**	**Parallel run**	**Impact analysis**	**Curiosity and peer review processes**	**Formal functional and component testing**	**User acceptance testing and hand-over documentation**	**External engagement and publications**	**RPI protocol process**	**Support structures and Service Level agreements**
Data flows sufficiently stable and reliable	✔									✔
Data Quality Assured and Impact Assessed	✔			✔	✔	✔				
Methods and requirements are fit for purpose		✔		✔	✔	✔		✔	✔	
Google Cloud Platform (GCP) is safe, secure and stable			✔				✔			✔
GCP pipelines produce expected indices and results within expected timeframes			✔	✔		✔				
The end-to-end process (from ingest to publication) can be run within the timelines of a regular monthly round			✔		✔		✔			
All expected publications and downstream outputs (new and existing) are clearly defined, of expected quality and can be produced in line with the regular monthly cycle			✔				✔			
Stakeholders are aware and supportive of intended changes				✔				✔	✔	

Scope of the parallel run

8. **Purpose**: The primary purposes of the parallel run are:

    a. In the early months of the parallel run: Ensuring the stability and accessibility of newly created pipelines, apps and related processes; i.e. ensuring that these processes can be reliably run within the timelines of a regular production round

    b. In the later months of the parallel run: Ensuring the stability, accessibility and completeness of the end-to-end process; including ensuring:

i. that all necessary publications and other outputs can be produced with the resources and time available,
ii. that all new and existing systems, data and processes interact seamlessly and reliably to produce our headline measures of inflation,
iii. that an effective handover of new processes to the existing teams is put in place.

9. **Timing**: We are currently planning to commence the parallel run in mid-July 2024, which would enable us to have 7 months of parallel running before moving these indices into production. This is referred to as a "parallel" run as it is expected to be produced in parallel (in terms of monthly timing) to the regular production cycle, with curiosity sessions held on briefing week to compare new aggregated indices to the traditional indices being prepared for publication.

10. **Processes**: In addition to new systems, data and methodologies, the parallel run will also incorporate new processes that are inherent to using groceries scanner data in production:
    a. Running new pipelines that stage data and produce the necessary indices and related outputs
    b. Weekly labelling processes for classification (the process by which products are classified to a given ONS product category) and relaunches (the process of linking a new product which is comparable to an existing product but just has a small change, for example, in packaging or size to deal with e.g. shrinkflation);
    c. Using new dashboards to validate data and indices produced;
    d. Working closely with different retailers and resolving challenges if/where necessary against production timelines; and
    e. Ensuring indices for traditional data still link reliably into downstream processes; incl. producing supplementary analysis currently produced using the Ingres system.
    f. Other processes inherent to using a new systems infrastructure such as security measures, processes for adding/removing users and approach to continuous improvements.

Scope of the impact analyses

11. We will also produce a range of impact analyses. This is a distinct process from the parallel run since impact analyses can be run at any point in time and are not necessarily dependent on the completion of the parallel run (although the final impact analysis results should align to figures produced during the parallel run).

12. Currently we plan for the impact analysis to cover at a minimum data from the start of 2020 to the end of 2023. There are two primary factors that will affect the start and end dates. Firstly, we are still working with individual retailers to see if we can have certain cleaned data re-delivered to extend the start date to 2019. Secondly, historic scanner data is currently only fully labelled up to the middle of 2022, so labelling all data up to the current date will require an extensive labelling process in the coming months. We will endeavour to conclude all

historic labelling by the time of the production of the impact analysis, but given resource constraints we still need to assess whether we will have been able to re-label all data by that time.

Impact analysis process

13. The rest of this section sets out the planned scope for the impact analysis to be produced, comparing outputs from the existing systems with those produced on the new platform. It details the proposed changes from the current system and how the impact of these changes will be assessed.

14. As we are assessing the impact of potential changes in a multi-step process, changes made earlier into the process will alter the input and therefore output in later steps. As there are multiple steps with multiple options, we will not aim to run every possible configuration of these options, as this would greatly increase the time, complexity, and challenge in interpreting and communicating the results. The approach we plan to take is therefore to run the pipeline end-to-end with current settings, then sequentially change each single option to assess its isolated impact; in a number of "steps" detailed below.

15. **Step 1: Setting a baseline through appropriate treatment of imputed COVID items:** During the pandemic some offline changes were made in spreadsheets to account for the unprecedented levels of items that needed imputing using the headline rate of inflation. Those calculations have not been in scope of new requirements so for the impact analysis we need to create a new baseline, which can be compared to published indices to understand the impact of not including these offline imputations, so this impact isn't conflated other methodological and data changes.

16. **Step 2: Impact of improved imputation methods for CPI & CPIH** This step would measure the impact of improved imputation methods as outlined in our 2021 article. These improvements would be made for CPIH and CPI only given the different classification hierarchy for RPI.

17. **Step 3: Impact of changing the lag in imputation.** Currently if a quote is given an N (non-comparable) code in April (for example), the quote would not contribute to the April or May index. The May index would then be used to impute a new base price for the product, which would then start contributing again in June. This step would measure the additional impact of changing to a 1 month lag, as is already done in central spreadsheets and for local collections where a new provider has been recruited.

18. **Step 4: Identifying a method for shop-type stratification**: This step would additionally assess the impact of different shop-type stratification options (listed below). This is a key methodological choice still to be made in the calculation of retailer weights, as set out in 2021 publication. The options that will be tested will be:
    a. Multiple/ independents (current method).
    b. Other big/ other small – an alternative method of weighting the local collection retailers according to market share

c. Implicit weighting – an alternative method of weighting the local collection retailers implicitly based on the number of quotes collected at each retailer rather than any explicit weight

19. **Step 5: Identifying the impact of introducing consumption segments:** This would build on the previous scenarios to understand the impact of introducing consumption segments and forcing stratification options on some items. This involves comparing two options:
    a. The current 1:1 mapping – all items would map to a single consumption segment (for example, basmati rice mapping to basmati rice). Note that we cannot incorporate grocery scanner data under this method, so it is primarily for illustration relative to the proposed N:1 mapping.
    b. Proposed N:1 mapping (as set out in our 2021 publication) – in some situations, there may be more than one item mapping to the same consumption segment (example, basmati and microwave rice mapping to "Rice"). This forces some additional stratification logic onto the item level, for example if basmati is not collected regionally but microwave rice is, then we would have to set the Rice consumption segment to also not be stratified regionally (and group together all regional microwave rice price quotes) otherwise we wouldn't have a consistent aggregation structure at the item level to aggregate up to consumption segment level.

20. **Step 6: Identifying the impact of introducing grocery scanner data.** This step would understand the impact of introducing grocery scanner data on top of all the earlier changes mentioned and hence (when compared to published estimates) would represent the full impact of all proposed changes. Specifically, this option involves removing all price quotes for strata that will be covered by groceries indices (for example, if we captured blueberries from Retailer A in our scanner data indices, we would remove all blueberries price quotes from Retailer A from our local collection data to avoid duplication). The indices from ADS and local collection would be aggregated together using market shares for the retailers using the selected methods as alluded to above.

Dawid Pienaar
Prices Division, Office for National Statistics
January 2024