

Constructing variance estimates for the UK Consumer Price Indices:

Progress update July 2025

Status: Work in progress

Expected publication: Alongside minutes

Purpose

1. The UK Statistics Authority, in approving the re-designation of CPIH as National Statistics (now “accredited official statistics”), included a requirement to explore and publish estimates of quality. Since then, work to assess the sampling variance of the consumer price indices has been underway.
2. This paper summarises the progress made since April 2024.

Actions

3. Members of the Panel are invited to:
 - note the progress with variance estimation
 - comment on the expectations of the impact of chainlinking
 - comment on the framework for variance estimation to be investigated to provide quality measures for indices constructed from scanner data and multilateral indices.

Details

4. Since the previous update in April 2024 ([APCP-T \(24\)04](#)) we have:
 - Completed the calculation of the variances due to price sampling using the ultimate cluster approach, over a long span. This produces variances which are slightly larger than expected, which may in part be due to the way the prices are chained together rather than chaining the higher level indices. In order to have a better view of this, yet a further round of calculation replicating the CPI chainlinking approach inside the variance calculation may be needed.
 - Considered the properties of the jackknife variance estimators of the elementary indices in design- and model-based situations. This work suggests that the jackknife estimator of the variance is the same under these different paradigms in both element and cluster sampling, which is helpful in providing a robust framework for the statistical interpretation of the variances. It also answers one of the questions about the difference between the design-based and model-based variances in the elementary indices, though further work is needed to extend this to aggregated indices.
 - Begun consideration of the approach to variance estimation with scanner data (and similar data sources) and multilateral indices.
5. When scanner data and similar sources are incorporated into the CPI, the index segments which are based on them will have no sampling variance in the classical sense, because the data will

essentially be completely observed. However, since there is no objective target parameter, there will still be uncertainty in the calculation of a price index. This may manifest in one of two counteracting ways – first, many more prices will be used in the calculations, so the variability of any parameter based on averaging will tend to be rather small, and second, a much wider range of products, including much rare ones, will be included, which may be expected to make the variability of the prices much larger, though it is not clear what impact it will have on the variability of the *changes* in the prices. The only framework that offers a reasonable approach is the model-based framework, and here we need to work with an appropriate model against which to evaluate the variability of the observed prices.

6. We continue to follow the plans for development of variance estimates. Some additional components are now in need of development:
 - The work done by Jim O'Donoghue at the beginning of investigations of the variance of the CPI is now very out of date, and does not adequately represent the variability in the index due to the current basket and weights. The next step is to make this into a repeatable calculation which can be applied to the LCF in any year, in order to produce variances of the weights appropriate to the index. There are some challenges with how to deal with the effect of input-output balancing, which is a component of the weighting process, and how to deal with weights sources which are not derived from the LCF.

Paul A. Smith, University of Southampton
July 2025