

Compendium

Limitations

Contact:
Philip Wales
macro@ons.gsi.gov.uk

Release date:
15 December 2014

Next release:
To be announced

Table of contents

1. [Section 8. Limitations](#)
2. [8.1 Common price indices](#)
3. [8.2 Aligning micro-level expenditure data and CPI weights](#)
4. [8.3 Data sources](#)
5. [Background notes](#)

1 . Section 8. Limitations

While the calculation of inflation rates for subsets of the household population is analytically straightforward, a range of data constraints make their estimation challenging in practice. As a result, this paper and others in the field make a range of simplifying assumptions. This section discusses these assumptions and the limits of our analysis with the aim of identifying areas for future work.

2 . 8.1 Common price indices

Perhaps the most important limitation of this analysis is the use of national price indices alongside sub-group specific expenditure weights. As set out in Section 2, an analysis of sub-group specific inflation rates would ideally use price indices and expenditure weights specific to each household¹. While the expenditure weights used here capture differences in the consumption patterns of different households, the lack of sub-group specific price indices means that this and other papers assume that all households face the same prices.

While this may be a fair assumption for some items – TV licences for instance, for which there is little variation in price – it is less likely to hold in product categories which comprise large numbers of heterogeneous items – such as second-hand cars or photographic, cinematographic & optical equipment. In these categories, the products included in the Consumer Prices Index (CPI) are selected to be representative of the purchases of all households and therefore capture ‘average’ price movements. As a result, they may be more or less representative of the prices that sub-groups experience.

The impact of this assumption on our analysis depends on the extent to which households face different prices for goods in the same COICOP class and on the extent to which within-class prices move together. As is summarised in Table 8.1 below, when different types of households purchase broadly similar types of goods within each COICOP² class and when intra-class price movements are similar, the impact of our assumption should be relatively small (top left quadrant). The greatest impact of the assumption arises where households are highly differentiated in their within-class purchases, and when the prices of those products move in very different directions (bottom right).

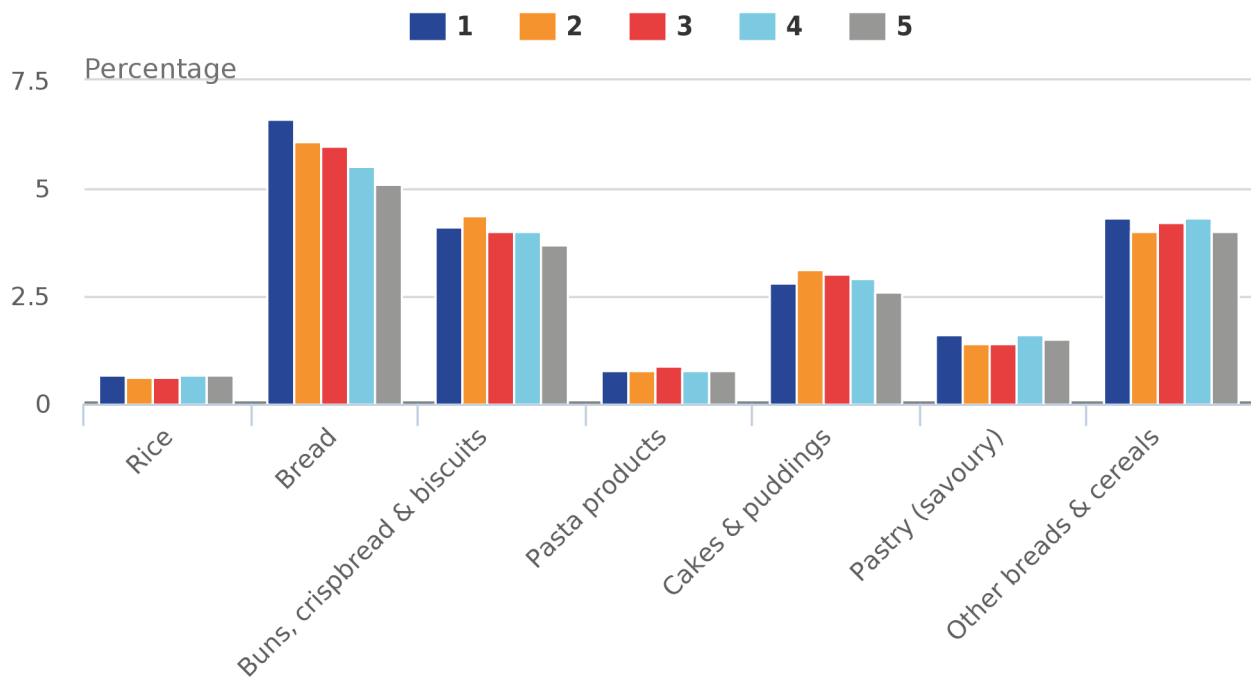
Table 8.1: Impact of intra-class price and product differentiation on sub-group inflation estimates

Degree of intra-class price integration	Degree of between household variation in within-class product purchases	
	Households buy similar within-class products	Households buy different within-class products
Intra-class product prices move at a similar rate	Low	Moderate-High
Intra-class product prices move at different rates	Low-Moderate	High

Source: Office for National Statistics

As data on the degree to which prices vary for different types of household are not available, it is not possible to quantify the impact of this limitation with any precision. Academic work on this topic is also limited: two studies based on households in the United States – Broda, Leibtag and Weinstein (2009), and Aguiar and Hurst (2007) – suggest that there is evidence of differential pricing for different household types, but as yet there is no comparable evidence for the UK. However, Figures 8.1 and 8.2 give some broad sense of its potential importance. Figure 8.1 shows the proportion of household food expenditure that is accounted for by sub-components of bread & cereals by equivalised income quintile in 2012. It shows that households vary in the type of within-class products that they purchase, with households in lower income quintiles allocating a greater share of their expenditure to bread in particular, than higher-income households. As these within-class and between-household differences in expenditure patterns are likely to vary in magnitude with the nature of the product considered, Figure 8.1 suggests that future work could usefully document these differences and take account of them in the calculation of household-level inflation rates.

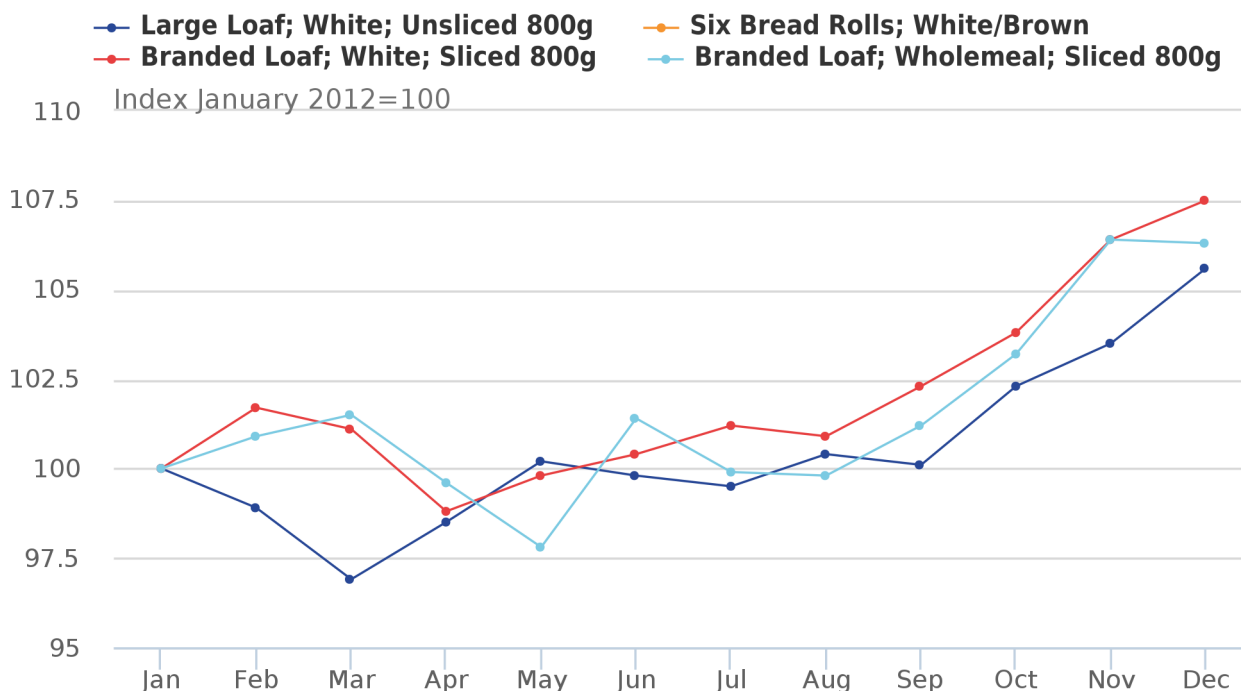
Figure 8.1: Food expenditure share allocated to sub-components of bread & cereals COICOP class by equivalised income quintile, %, 2012



Source: Living Costs and Food Survey - Office for National Statistics

If there is some evidence of within-class differences in expenditure patterns, to what extent do the prices for these different products vary through time? Figure 8.2 explores this question using the micro-level price quote data on which the CPI is constructed. It shows the distribution of price growth observed for four different types of bread in 2012. The price of branded white loaves of bread for instance, has increased by around 7.5% over this period, while the price of six bread rolls increased by just 2.3%. If these differences are replicated in different classes and over different time periods, then the impact of our assumption may be relatively large.

Figure 8.2: Prices of selected 'bread' products: January 2012=100



Source: Office for National Statistics

Taken together, Figures 8.1 and 8.2 suggest that there is potential for the results presented in this paper to be affected by within-class differences in expenditure and price growth. However, without more detailed information on both expenditure and prices, it is not possible to estimate either the direction or the size of this effect. This challenge we leave to future work.

Notes for 8.1 Common price indices

1. See Section 3 for more details.
2. COICOP is the Classification of Individual Consumption According to Purpose. See Section 2 and Appendix B for more detail.

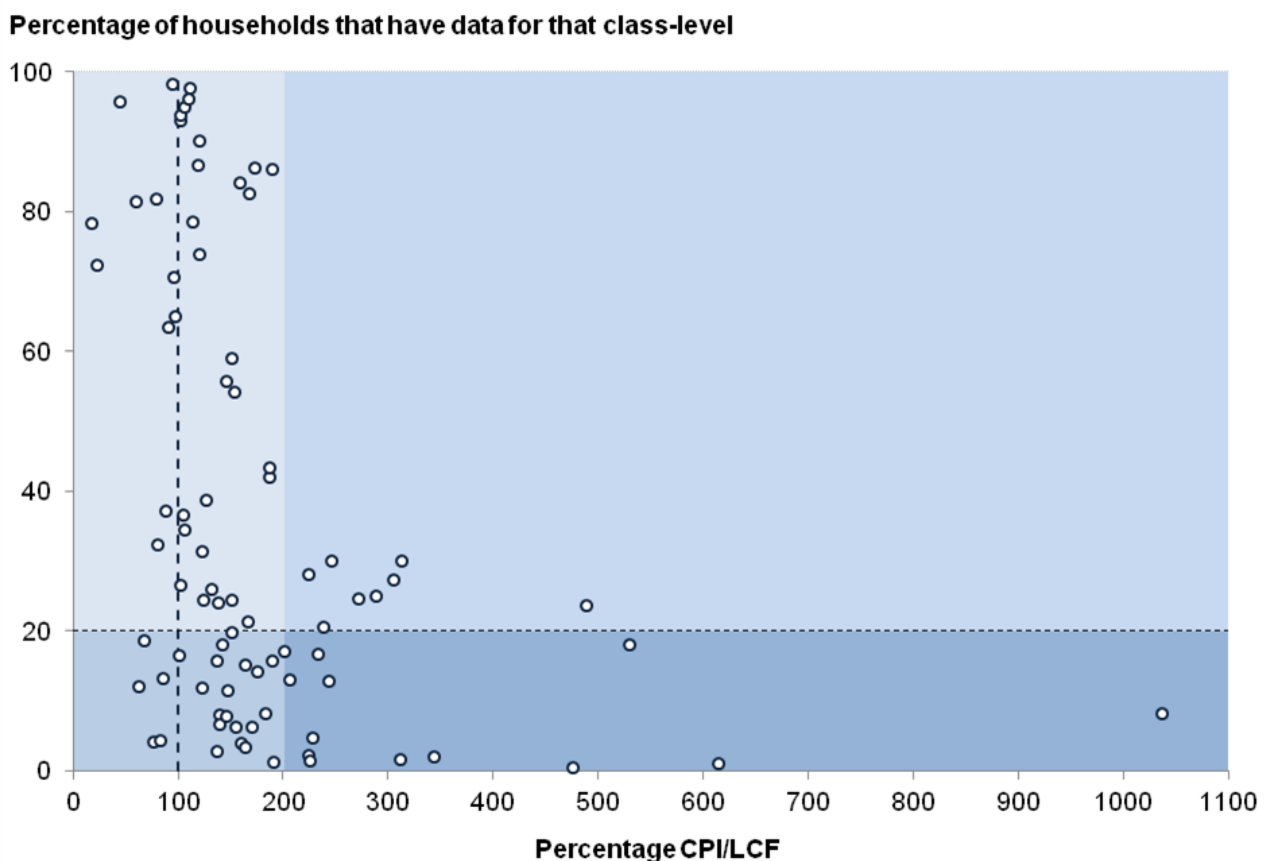
3 . 8.2 Aligning micro-level expenditure data and CPI weights

A second set of assumptions used in this paper concerns the methods that are used to harmonise the micro-level expenditure data from the Living Costs and Food Survey (LCF) and the expenditure weights that are used in the construction of the Consumer Prices Index. As set out in Section 3.1 above, these may differ for a range of reasons, but in particular because the LCF is not the sole source of expenditure data in the CPI. The broad method used to harmonise these two sets of expenditure data is set out in Section 3.2, and involves the allocation of the CPI expenditure totals to observed LCF households in proportion to the spending that they report in the LCF.

Harmonising the CPI expenditure weights with the micro-level data from the LCF in this way requires several assumptions. First, we assume that the distribution of household-level spending on each COICOP class reported in the LCF is representative of the 'true' underlying distribution. This, apparently anodyne assumption has two important corollaries: (a) that households do not censor their spending patterns in the LCF; and (b) that all types of households are equally likely to over- or under-state expenditure on a given product type. To see the importance of these assumptions, consider spending on alcohol and tobacco, both of which are thought to be under-recorded by the LCF (ONS, 2014d). The first corollary states that no household reports zero expenditure on these categories if they have strictly positive expenditure on alcohol and tobacco during the reporting period. The second corollary implies that – having met this first condition – no household type is more or less likely to over- or under-report spending on alcohol or tobacco. Both assumptions are relatively strong, and unfortunately there is no way to estimate the impact of these biases without additional data¹.

Secondly, where the expenditure weight in the CPI is based on data other than the household survey, the differences between the LCF and CPI expenditure totals can be large (see Figure 3.3). To assess the effect of allocating these new expenditure totals to the observed household units, we conducted a set of plausibility checks. The first of these is presented in Figure 8.3 below. The horizontal axis shows the difference between the LCF and the CPI expenditure total as a proportion of the LCF total: numbers close to 100 therefore indicate close correspondence between the two weights, while observations further from 100 indicate greater differences. The vertical axis plots the number of households who report strictly positive expenditure. Each point is a single COICOP class, representing expenditure on a given set of products. The chart is divided into four shaded 'zones'. In the top left are COICOP classes in which the LCF and CPI expenditure totals are of broadly similar orders of magnitude (defined as CPI expenditure less than double the LCF total) and where the number of households reporting strictly positive expenditure is relatively high (defined as above 20%). In this segment are many products for which the LCF is taken as the basis for the CPI weights. Immediately below this quadrant are instances where the number of households reporting strictly positive expenditure is relatively small, but where the CPI and LCF expenditure totals are fairly similar. Points in these two quadrants will introduce the least potential bias.

Figure 8.3: Difference between LCF and CPI as a % of LCF and % Households reporting positive expenditure, average 2002-2014



Source: ONS Calculations

Notes:

1. Figure is a scatter plot showing the 85 class-level categories of the CPI. The vertical axis plots the percentage of surveyed households who report some expenditure on that class-level item. The horizontal axis shows the percentage difference between the expenditure total in the CPI and LCF.
2. The Figure excludes class-level category 06.3.1: Medical & paramedic services as the CPI expenditure total is around 230 times bigger than in the LCF.

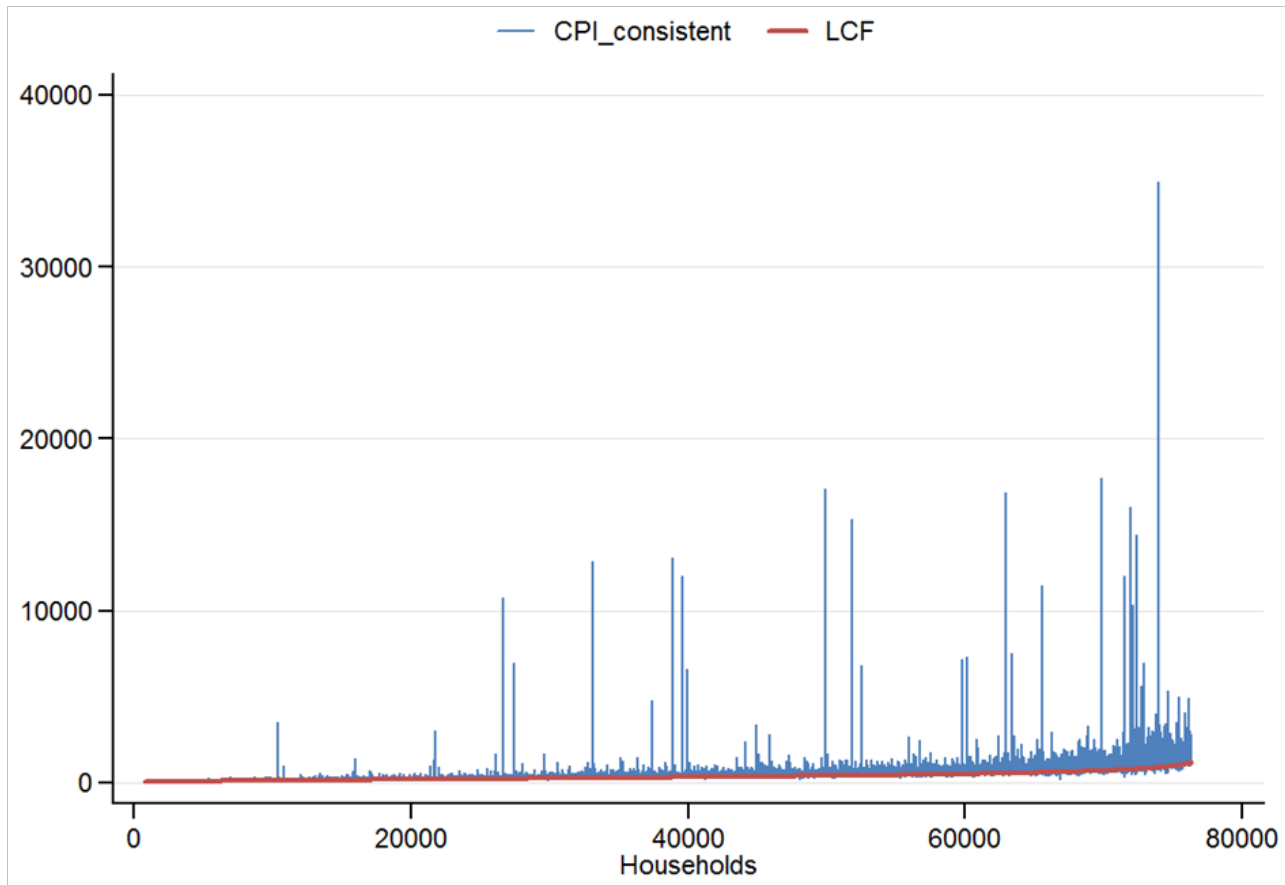
Points in the top right hand quadrant, by contrast, represent classes where CPI expenditure is high relative to the LCF total, and where a relatively large number of households have reported positive spending. In these cases, the potential for bias is also limited, as a large proportion of households will be affected by the micro-level attribution mechanism. However, it is points in the bottom right-hand quadrant that present the most difficulty: these are COICOP classes in which the CPI expenditure total is more than double the LCF total, and in which fewer than 20% of households report spending. Medical & paramedic services remains an outlier, with just 66 households reporting spending on this COICOP class-level over the eleven years of available data.

In these cases, and where the proportion of LCF respondents reporting expenditure is relatively small, we adjust our methodology to avoid perverse results. Implementing this adjustment requires the assumption that it is appropriate to allocate total spending on a COICOP class (6.3.1 Medical & paramedic services, for instance) using reported household expenditure on a higher aggregate – (6 Health, for instance). This assumption ensures that our methodology does not allocate very high levels of spending to a relatively small number of households, which in turn would distort the picture of household inflation.

To examine the impact of this assumption on our analysis, we conducted two further tests. First, we inspected the distribution of spending across different equivalised income decile groups in both our class-level categories and the division-level proxies. This reassured us that the process does not reallocate spending from one income group to another. Secondly, we altered the thresholds – set out in Figure 8.3 – to test the sensitivity of our findings to changes in our methods. The results of these sensitivity tests are presented in Appendix C. These indicate that changes to the reporting thresholds set in Figure 8.3 have little impact on the overall profile of inflation rates or the headline results of this work.

How do these assumptions affect total household spending? Figure 8.4 presents a final plausibility test that ONS carried out to test the micro-level attribution mechanism. Figure 8.4 ranks all included households by their expenditure in the LCF, shown in the red line. Overlaid on this is the household spending total following the micro-level reattribution of the CPI expenditure weights, shown in blue. Panel A shows the results of the reattribution without implementing an additional rule for the difficult cases identified in Figure 8.3 above, while Panel B shows the final results after the implementation of the additional rule.

Figure 8.4A: Total household expenditure, LCF and CPI-reconciled totals, simple attribution

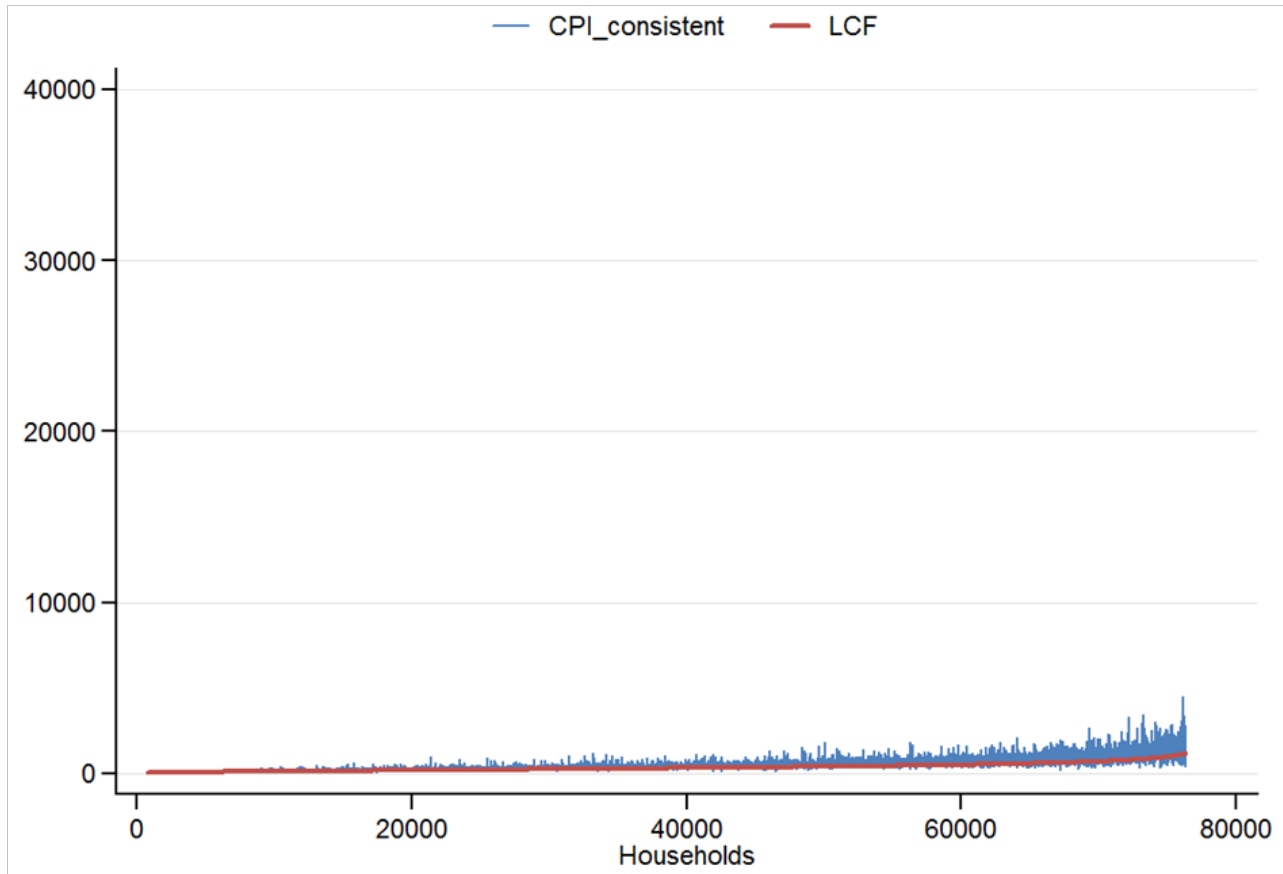


Source: ONS Calculations

Notes:

1. These figures show the LCF and imputed CPI-reported expenditure totals for each observed household.
2. Figure excludes the top and bottom 1% of households by LCF spending.

Figure 8.4B: Total household expenditure, LCF and CPI-reconciled totals, attribution including proxy categories



Source: ONS Calculations

Notes:

1. These figures show the LCF and imputed CPI-reported expenditure totals for each observed household.
2. Figure excludes the top and bottom 1% of households by LCF spending

Panel A suggests that without the additional rule, the simplest reattribution of the CPI weights adjusts total expenditure for some households considerably, as is shown by the 'spiky' blue line. In these cases, adjusted spending is substantially higher than the LCF total, which in turn points to a risk that our method is affecting the results. However, after the implementation of the additional rule in Panel B, the distribution of expenditure is broadly maintained, and very few households' expenditure totals are substantially distorted. These results, and the broad similarity between the results of the LCF and CPI-consistent analysis, give us confidence that our results are the product of well-defined trends in household experience, rather than our methods.

Finally, it is worth noting that this alignment process delivers a micro-level expenditure dataset that is consistent with the CPI weights, not with the National Accounts. The difference here arises because – as discussed in Section 3 above – the CPI weights are price up-dated from a previous year and are not revised. Implicit in this mechanism is the assumption that households do not substitute in the face of changing prices between the point of survey and the fixing of the expenditure weights. For our work, this implies that while the alignment process delivers the CPI weights exactly, it will not deliver the National Accounts expenditure totals.

Notes for 8.2 aligning micro-level expenditure data and CPI weights

1. In principle, it is possible to estimate 'imputed' values for households who do not report any expenditure in a given COICOP. This would be likely to involve a two-step model, first estimating the likelihood of positive expenditure on a given product, before estimating the extent of that expenditure. However, this approach involves a further set of explicit and implicit assumptions. Here, we have chosen to keep our approach simple and leave this line of inquiry for future research.

4 . 8.3 Data sources

There is also a range of additional limitations that relate to the data sources used in this paper, rather than the methods employed to calculate price indices and inflation rates. First, while the LCF is a relatively large, continuous survey of household expenditure, it places a burden on respondents. As a result, the response rate varies – ranging between 50% and 60%. As there are no obvious candidate variables that could be used as exclusion restrictions, we have not been able to model this process of non-response. This may affect our results if non-reporting households have very different patterns of expenditure to those who do report, although non-response weighting is used throughout to alleviate this issue. Analysis of the response rate suggests that some types of household are less likely to respond to the LCF (Bright et al., 2009), but without more detailed information it is difficult to assess the likely size or direction of this effect. Secondly, the LCF does not cover some types of household that might be of interest. In particular, it does not cover student halls and other communal establishments.

Thirdly, housing costs are a further area of weakness in our work. In particular, the lack of information required to produce micro-level estimates of rental equivalence, the ONS' preferred measure, means that the price indices calculated here – while consistent with the national CPI – exclude the costs associated with housing for a large number of households. This limitation is particularly concerning, because it is likely to have a differential impact across some of the sub-groups of interest. For instance, renters (whose housing costs are captured in this work) are more prevalent among low-income, low-expenditure and non-retired households than owner-occupiers (whose housing costs are less well captured), who are more likely to be in the high-income, high-expenditure and retired household groups. All of these issues we leave for future work.

5. Background notes

1. Details of the policy governing the release of new data are available by visiting www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html or from the Media Relations Office email: media.relations@ons.gsi.gov.uk