Office for National Statistics

Compendium

Methods, data and literature

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

Release date:
15 December 2014

Next release:
To be announced

Table of contents

1. Section 2. Theory

2. Section 3. Data & methods

3. Section 4. Context & literature

4. Background notes

1 . Section 2. Theory

A price index has two basic ingredients: data on the quantity of products purchased and information about the prices of those products. These two ingredients may be combined in various ways to produce different forms of price index: in the UK, the Consumer Prices Index (CPI) uses a Lowe price index, which is a Laspeyres-type index [1]. This uses expenditure data from the previous period alongside information on prices in this and previous periods, and is shown in equation [2.1]:

Equation 2.1

$$I_r^t = \frac{\sum_{i=1}^{n} p_i^t q_i^r}{\sum_{i=1}^{n} p_i^r q_i^r}$$

where I is the index value, p is the price level, q is the volume purchased and r, t and i index the reference period, time and items respectively. In more simple terms, this formulation involves using changes in prices alongside expenditure weights from a fixed period. The prices of items that account for a larger (smaller) fraction of expenditure in the reference period are given a greater (lesser) weight in the calculation of the overall index. From this perspective, the formulation of a price index for a subset of households is trivial. For a subset of households, A, the price index equivalent to [2.1] is calculated using data on the expenditure of those households and the prices which they face:

Equation 2.2

$$I_{A,r}^t = \frac{\sum_{i=1}^{n} p_{A,i}^t q_{A,i}^r}{\sum_{i=1}^{n} p_{A,i}^r q_{A,i}^r}$$

By extension, the equivalent price index for any given household, a, is given by [2.3]:

Equation 2.3

$$I_{a,r}^t = \frac{\sum_{i=1}^{n} p_{a,i}^t q_{a,i}^r}{\sum_{i=1}^{n} p_{a,i}^r q_{a,i}^r}$$

Equations [2.1] to [2.3] therefore set out the information that is needed to calculate price indices for all households, a subset of households and an individual household respectively. However, equations [2.2] and [2.3] also have the property that they if they are weighted to reflect the spending of the relevant unit, the all-household price index can be recovered:

Equation 2.4

$$I_r^t \equiv \sum_{A,r} \frac{e_A^t}{E^t} I_{A,r}^t \equiv \sum_{a,r} \frac{e_a^t}{E^t} I_{a,r}^t$$

where e and E are unit and whole-economy household expenditure respectively. This formulation highlights a further feature of these price indices that is relevant for this analysis. Equation [2.4] shows that the standard Laspeyres-type price index used in the CPI weights the price experience of different households by their share of expenditure. While this is not an explicit design of the methodology – which more heavily weights the prices of high-expenditure items – a corollary of this approach is that households that spend more each period have a greater weight in the calculation of the CPI than households who spend less [2].

This can lead the price experience of a subset of households to differ from the published CPI – in particular among those households whose expenditure patterns differ substantially from that of the average for the sector as a whole. Price indices of this form are described as having ‘plutocratic weights’, and have the feature that they more heavily weight high-spending households. However, while this is standard international practice, alternative weighting mechanisms can be deployed.

One potentially interesting alternative formulation is a so-called ‘democratic [3]’ price index, which is shown in equation [2.5], where n represents the number of households:

Equation 2.5

$$I_r^t \gtreqless I_r^{Dem,t} = \sum_{a,r} \frac{1}{n} . I_{a,r}^t$$

In this formulation, each household receives an equal weight, regardless of their level of spending. The aggregate democratic price index consequently takes the average of the values for each household. In the absence of longitudinal data, this form of price index uses expenditure weights calculated by simply averaging the weight assigned to each product across households.

Note that the level of aggregation in [2.5] is crucial: taking price indices at anything above the household-level (which would implicitly require weighting of some form), would place different weights on different households. As a result, a democratic index is a relatively data-intensive form of price index, requiring household-level expenditure and price information. In general, the difference between the plutocratic and democratic indices will be larger (smaller) when the composition of household spending varies considerably (very little) across households.

Notes for Section 2. Theory:

1. The CPI is a Lowe index, in the sense that it uses current-period price information alongside expenditure weights that are price-updated. This latter feature distinguishes it from a Laspeyres price index, which uses current period price information with observed, previous period expenditure weights. For notational simplicity, we present these price-updated weights as if they were observed and, as a consequence, our treatment here appears more like a Laspeyres index. See Appendix A.

2. To see this, consider an economy with two households: one high-spending and one low spending household. Suppose that the majority of the high spending household purchases are devoted to high-inflation products, while the majority of the low-spending household purchases are devoted to low-inflation products. The CPI for this economy – which uses the total amount of household spending – will more closely reflect the inflation experience of the high-spending household, as the weights for the sector as a whole are closer to its expenditure shares than the low-spending household. As a result, the CPI will be above the inflation rate experienced by the low-spending household, and close to (although below) the inflation experience of the high-spending household. The degree of the inflation differential will vary depending on the extent to which household spending shares differ.

3. Note that the naming convention here can be misleading: In a ‘democratic’ index, each household is given an equal weight, rather than each individual, which might be implied from its name. A ‘truly’ democratic index would weight each person in an economy equally, and would deviate from the popular convention of a ‘democratic’ index to the extent of variation in household size. Arguably, a still ‘truer’ index would use longitudinal data to observe movements in expenditure patterns for the same individuals through time; however, this approach is data-intensive, challenging to implement, and its interpretation not straightforward.

2 . Section 3. Data & methods

3.1 Data

As set out in Section 2, price indices have two ingredients: data on expenditure by product and information about prices for each of these products. This section sets out the data used in this paper to calculate aggregate, decile and sub-group price indices and rates of inflation.

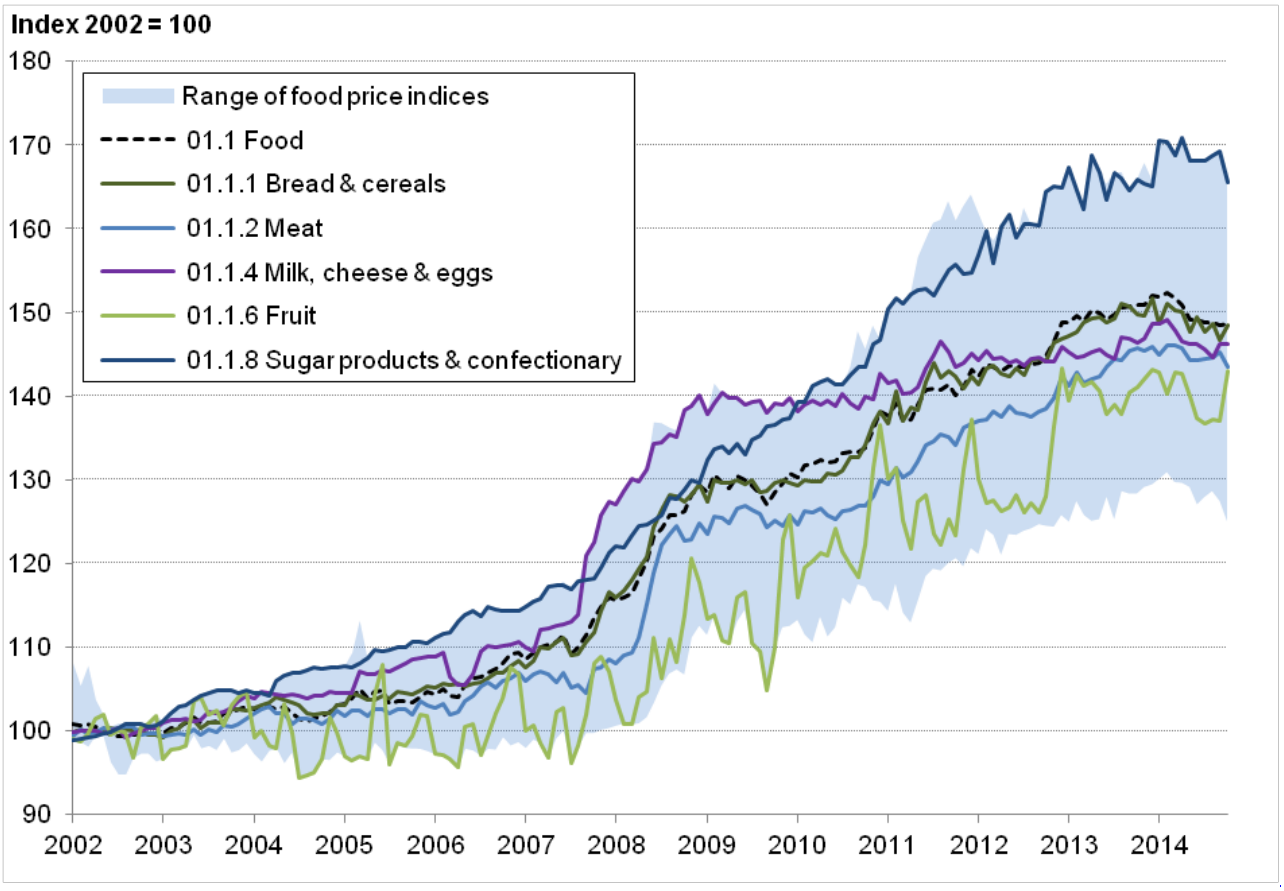
3.1.1 Price data

The price data that are used in this paper are taken from the Consumer Prices Index (CPI). The CPI is calculated from around 100,000 price quotes taken every month for a range of products in different shop types across the regions and nations of the UK. These are supplemented by a centralised collection of a further 80,000 monthly price quotes, ensuring that the CPI is based on around 180,000 price quotes for around 700 goods and services each month. These price quotes are weighted using expenditure data from the National Accounts to ensure that the basket of goods and services reflects the spending of the household sector as a whole (ONS, 2014a).

As household-level expenditure data for individual products can be volatile and intractable, this analysis uses expenditure and price data that is aggregated to the 85 class-level categories defined in the Classification of Individual Consumption According to Purpose (COICOP) (UNSD, 2014)[1]. This detailed dataset therefore provides information about how prices have evolved for 85 groups of goods and services, ranging from bread & cereal to pharmaceutical products, from health insurance to air travel products. These indices are used in their unrounded format, as they are produced prior to the publication process to ensure that errors arising from data aggregation are minimised.

Figure 3.1 shows price indices for several COICOP class-level categories that sit within the 'food' group for the 2002 to October 2014 period. It highlights five specific class-level price indices, as well as the weighted movement for the entire group (the dashed line) and summarises the movement of the remaining four price indices in this group using a swathe to denote the range. Figure 3.1 gives some idea of the detail of the categories on which this analysis is based, and demonstrates that there can be substantial differences in price movements across different COICOP classes. For instance, the prices of sugar products & confectionary grew by 65.6% over this period, compared with 43.0% and 43.6% for fruit products and meat products respectively [2].

Figure 3.1: Class-level price indices for 1.1: Food, 2002=100



Source: Office for National Statistics

However, the use of these data introduces the first of several limitations into our analysis. As shown in equation [2.3] above, the calculation of 'true' sub-group specific price indices requires the use of household-specific prices. However, as price data are collected from retailers rather than by asking households which prices they face, separate price indices are not available for different types of household. As a result and in common with previous, similar studies, this analysis assumes that households all face the same class-level CPI average prices. This limitation is discussed in greater detail in Section 8.

3.1.2 Expenditure data

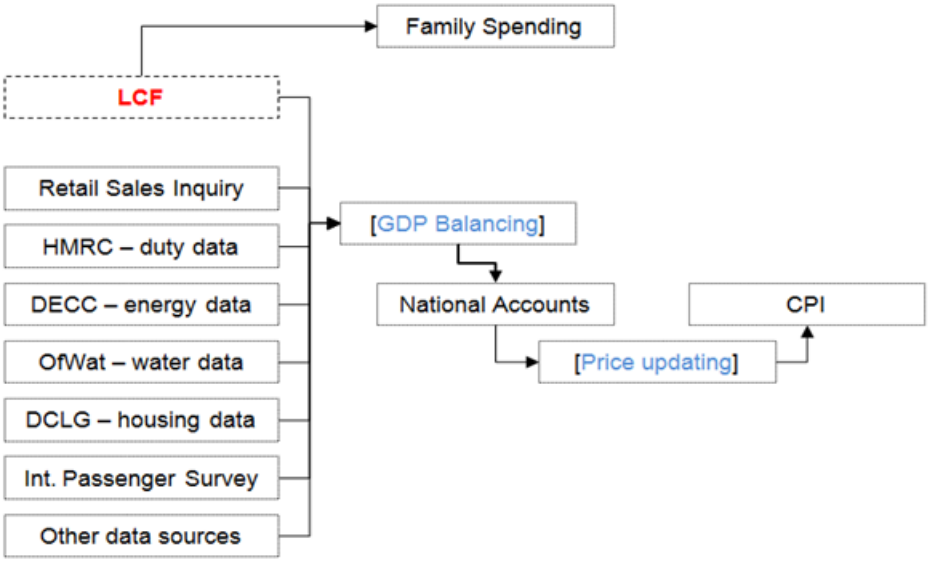
The expenditure data used in this analysis come from several different sources. First, household-level expenditure data are taken from the Living Costs and Food Survey (LCF). This continuous, cross-sectional survey gathers detailed expenditure data from between 5,500 and 6,500 households per year through a structured interview and an expenditure diary. A range of data, including the size and composition of the household, the household's tenure, the level of any mortgage interest payments and household income are also gathered, alongside the components required to calculate household-level expenditure for each of the 85 COICOP class-level categories. As a consequence, it allows a more detailed examination of spending by different household types than any other expenditure survey carried out by ONS.

The underlying LCF household-level sample consists of 73,506 households, surveyed between Q1 2002 and Q4 2013. A preliminary analysis of this sample suggested that there were a small number of households whose expenditure we regarded as implausibly concentrated on a single product type. We dropped 125 households (0.17% of the sample) on the basis that 80% or more of their total expenditure was concentrated in a single COICOP category. Secondly, we dropped a further 600 households (0.82% of the sample) who reported negative spending on any COICOP class – possibly reflecting the un-winding of prior overpayment. Taken together, these exclusions amount to 0.99% of the starting sample, and have no discernible impact on our results.

In addition to micro-level data from the LCF, this analysis also makes use of the aggregate household spending data which underpin the weights used in the construction of the CPI. Using these data allows us to (a) replicate the CPI directly; (b) calculate the difference between the published index and the price experience of households; and (c) analyse the impact of different weighting structures on price outcomes. These data were provided to us as annual expenditure totals, which we aggregated to expenditure totals for the class-level of COICOP.

How and why do the weights from the LCF and the CPI differ? Figure 3.2 shows a simplified process map for the calculation of CPI weights. While the LCF weights – as published in the 'Family Spending' release (ONS, 2014c) – are an input for the National Accounts and therefore for the CPI weights, it is only one of a number of sources used to estimate household expenditure. Alternative sources are used where the LCF is believed to under-report expenditure (including Alcohol and Tobacco) (ONS, 2012), where data quality is deemed to be stronger from administrative sources (including Energy) (ONS, 2014d), or where the concepts captured in the National Accounts differ from the pure expenditure estimates collected in the LCF (ONS, 2014a). This third case applies in particular to the costs of insurance (which are collected on a gross payments basis in the LCF, but on a net payments basis – after insurance payouts – in the National Accounts), used car purchases (collected on a gross expenditure basis, but measured as net household acquisitions in the National Accounts (ONS, 2014e) and estimates of the costs of financial services.

Figure 3.2: Input data for the calculation of the CPI weights

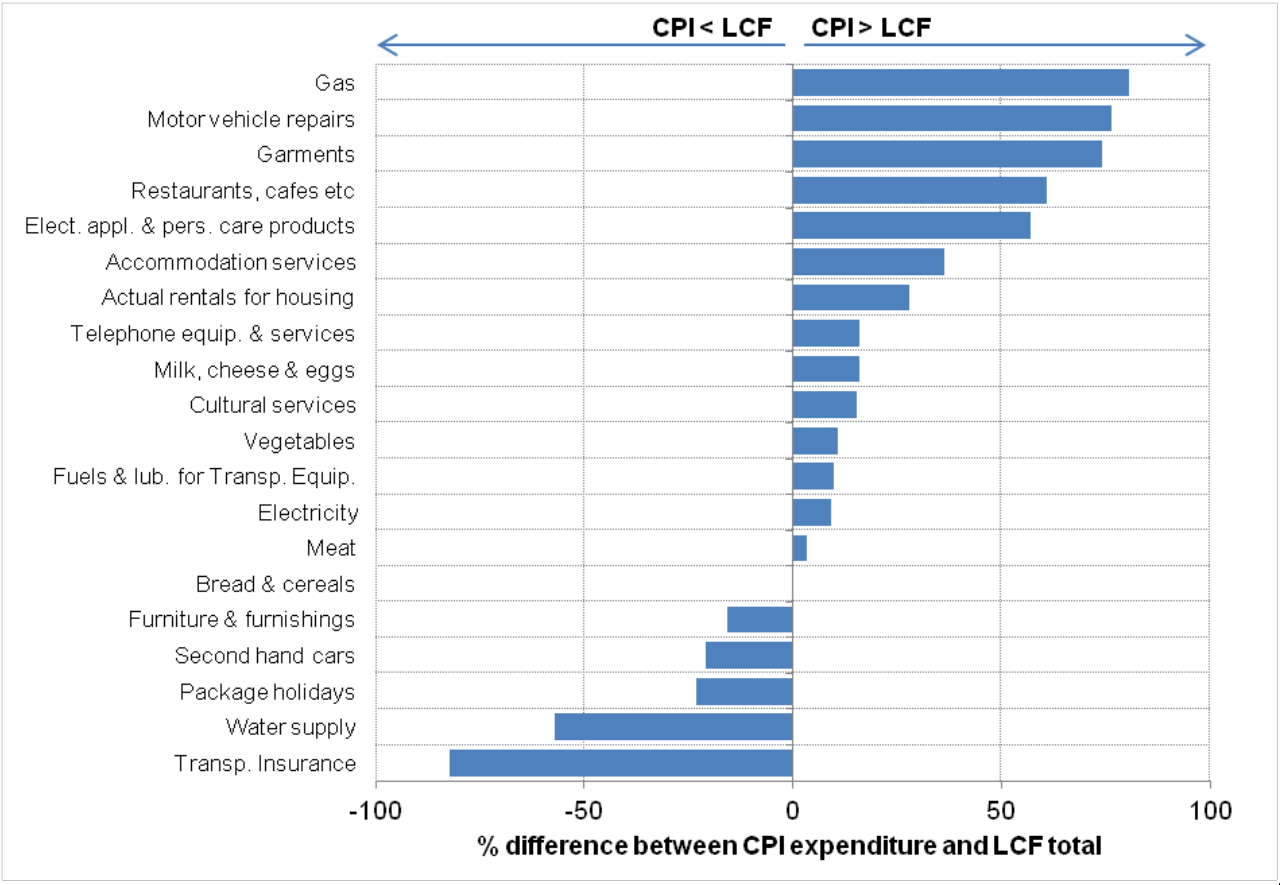


Notes:

1. Figure shows a number of the sources and processes used in the compilation of the CPI. LCF is the Living Costs and Food Survey, HMRC is Her Majesty's Revenue and Customs, DECC is the Department of Energy and Climate Change, OfWat is the water regulator, DCLG is the Department for Communities and Local Government, Int. Passenger Survey is the International Passenger Survey.

Secondly, expenditure totals from the LCF and the National Accounts may differ for a range of data processing reasons (shown in blue in Figure 3.2). Estimates of expenditure from the LCF may be affected by the GDP balancing process before they are used to calculate the weights for CPI. Additional adjustments are made to account for the spending of households from overseas in the UK. Further differences also arise because of issues of timing: to calculate current-year CPI expenditure weights each January (when LCF data for the current year is not yet available), observed expenditure from a previous year is 'price updated' (see Appendix A). This involves taking the expenditure totals in this previous year and imputing their current value using recent price changes. All of these practices are common across countries, but result in differences between the equivalent CPI and LCF estimates of household spending, shown for the 20 highest class-level expenditure categories in the LCF in Figure 3.3.

Figure 3.3: Difference between CPI and LCF expenditure totals in selected COICOP classes, %, 2011



Source: ONS Calculations

Notes:

1. This figure shows the difference between total LCF and CPI expenditure as a fraction of LCF expenditure for each product type in 2011. Bars to the right (left) of the axis indicate that the expenditure total is larger (smaller) in the CPI.

As the preceding discussion implies, Figure 3.3 suggests that the differences between the LCF and CPI weights are largest where alternative sources are used, or where the measurement concepts differ between the CPI and LCF. Household spending on gas is higher in the CPI than in the LCF as expenditure in the former is based on information from the Department for Energy and Climate Change, which is notably higher than the LCF estimates. Equally, the weight accorded to transport insurance in the LCF – which captures the cost of insurance premiums – is notably lower than the CPI estimate – which captures the cost of premiums less any claims.

A more detailed discussion of these differences and their impact on this analysis is deferred to the following section. However, it is worth noting at this point that a natural consequence of this discussion is that a price index based on LCF expenditure weights alone will not recover the CPI rate of inflation. Only when these alternative data sources are used and measurement concepts aligned are the CPI weights and inflation rate recoverable.

3.2 Methodology

The methods used in this analysis mirror those used in the calculation of the CPI. Unrounded class-level price indices for each month are taken from the CPI and placed alongside appropriate expenditure weights to produce an aggregate price index. The resulting indices are double chain-linked; first in January, which accounts for the annual changes in the COICOP weights for the class-, group- and division-level products (as set out in Appendix B). A further chaining step, to account for changes in the basket of representative items – the goods and services that are aggregated up to form the class-level of CPI – occurs in February [3]. To calculate the annual inflation rates, the monthly observations for each group are averaged across the year, and rates of change are estimated [4].

As a result, the only singular element of this work is in the construction of the expenditure weights, for which there are several candidate sources of data. To ensure the robustness of our analysis and to present interesting differences between measures, this paper replicates all of its results using three different sets of household spending data. The differences between the three resulting sets of expenditure weights are set out below. In the results that follow in Section 5, we focus on the third of these datasets. Section 6 uses the second dataset to assess the inclusion of housing costs on inflation rates. The full results using each set of weights are available in the Reference Tables.

Dataset 1: Weighted expenditure from the LCF

Dataset 1: Weighted expenditure from the LCF (74.5 Kb Excel sheet)

The first set of expenditure weights used in this analysis is the most straightforward. Using the micro-level data from the LCF, we calculate estimates of spending in each of the 85 class-level COICOP categories for each surveyed household. These totals are weighted to reflect the population as a whole and then aggregated across various sub-groups to yield sub-group specific expenditure weights. More explicitly, we aggregate these household-level weighted expenditure totals into: (a) equivalised [5] disposable income deciles; (b) equivalised expenditure deciles; (c) households with and without children; and (d) retired and non-retired households [6]. This is repeated for each year of our data, yielding expenditure weights for each sub-group in each period. These weights are used alongside the class-level COICOP price indices from the CPI. The resulting series are aggregated using the same process as for the CPI [7] and then averaged across the year.

Dataset 2: Weighted expenditure from the LCF and Mortgage Interest Payments

Dataset 2: Weighted expenditure from the LCF and Mortgage Interest Payments (74.5 Kb Excel sheet)

While the 85 class-level categories of COICOP include a range of different types of expenditure, they exclude any costs associated with the owner occupation of dwellings. Changes in rental costs, by contrast, are used in the CPI and are included in the 85 class-level categories. While the precise mechanism by which housing costs are included in price indices is a matter of some debate (ILO, 2004), the inclusion of housing costs for some of the population (those who rent) and not for others (home-owners with mortgages or owner occupiers) is a short-coming of our work, in particular as different forms of tenure will be more prevalent in some sub-groups than others.

In the context of our work, there are several different ways that housing costs for non-renters could be incorporated, some more difficult than others. In particular, perhaps the most attractive avenue is to produce a measure of sub-group inflation consistent with CPIH, including changes in the cost of owner occupation through the calculation of rental equivalence (ONS, 2014a). However, the production of micro-level estimates of rental equivalence is highly data intensive, requiring a complex array of data on different forms of tenure, housing and geographical area. This has been left for future work.

However, the costs of owner occupation – and in particular the costs associated with mortgage repayments – are often non-trivial fractions of household budgets. As a consequence, we create a second dataset, in which the 85 COICOP categories from the LCF are supplemented by interest payments on mortgages (excluding capital repayments). The resulting weights are used alongside the same CPI price indices as above, as well as the mortgage interest price index from the Retail Prices Index. While we recognise that this is a partial measure of housing costs – and in particular fails to capture the costs associated with owner occupation for households who do not have mortgages – it does allow us to consider a broader measure of housing costs than in Dataset 1. It is also worth noting that the treatment of insurance and the purchase of second hand vehicles remains on the ‘gross’ basis here – excluding the value of any claims or inter-household transfers respectively.

Dataset 3: Reconciling National Accounts totals with the LCF micro-data

Dataset 3: Reconciling National Accounts totals with the LCF micro-data (74.5 Kb Excel sheet)

The third expenditure dataset – on which the majority of the analysis in this paper is based – involves a micro-level reconciliation of the LCF and CPI expenditure weights, which differ for a range of reasons (see Section 3.1). This dataset – which represents one of this paper’s primary contributions to the literature – is composed of household-level expenditure estimates which aggregate to the CPI expenditure weights.

In order to produce a dataset of household-level spending estimates that is consistent with the CPI, this paper makes a series of assumptions designed to allocate the CPI expenditure total across the observed LCF households. In principle, there are several different ways that this could be achieved. Under our method, we seek to impose as few assumptions as possible on the data, and consequently employ a relatively simple rule which divides reported total CPI expenditure on each COICOP class among the households we observe in the LCF in proportion to their observed spending on that class-level category:

Equation 3.1

$$e^{CPI}_{a,i,t} = \frac{e^{LCF}_{a,i,t}}{\sum_{a} e^{LCF}_{a,i,t}} . e^{CPI}_{i,t}$$

where e is the level of expenditure consistent either with the CPI or LCF and where a, i and t index households, COICOP classes and time respectively. More simply, equation [3.1] states that total CPI-consistent spending on a given product is divided among the observed households in proportion to the share of total observed spending on that product reported in the LCF. Households that report more (less) expenditure on a given product are awarded a greater (lesser) fraction of total expenditure taken from the CPI. For instance, if an observed household [8] accounts for 0.05% of total purchases of bread & cereal products in the LCF, it is allocated the same fraction of the CPI expenditure total on bread & cereal. This simple attribution mechanism is a second limitation of our analysis – discussed in greater detail in Section 8 – and requires an important assumption: that where there are differences between the LCF and National Accounts totals for a given COICOP, these difference arise because all households over- or under-report their expenditure by the same proportion.

Extending the dataset to 2014

While the CPI weights and price indices that are used in this paper are available from December 2000 to October 2014, the LCF expenditure data that we use is only available on a comparable basis for the period 2002 to 2013. To extend our work to the most recent data, we price update the expenditure weights for individual households observed in 2013 to yield an imputed household-level dataset for 2014. The aggregate CPI weight for 2014 is consequently distributed among these imputed households in a manner discussed above, and in greater length in Section 8. This approach is consistent with the calculation of the weights for the CPI itself, and will only affect our results if different sub-groups experience stronger or weaker substitution effects (Levell and Oldfield, 2011). This assumption adds a further 5,117 imputed household observations to our dataset, resulting in a total sample size of 77,898 across the 2002 to 2014 period. Finally, note that while this approach allows us to extend our analysis by equivalised expenditure deciles and household types to 2014, the income data required to complete an analysis by income decile only extends to 2013. Consequently, our analysis presents inflation rates for the former groups of households to 2014, but only to 2013 for households in different income deciles.

Notes for Section 3. Data & methods

1. See Appendix B for more information

2. These differences make a strong case for using more disaggregated expenditure to analyse sub-group specific inflation rates: higher level price indices will only reflect the price experience of sub-groups if all households purchase the group's class-level components in equal proportions. Further discussion of the appropriate level of aggregation is deferred to the following section and Section 8.

3. Further technical information on how the CPI is constructed is available in the CPI Technical Manual (ONS, 2014a), while a higher level summary is available as an infographic (ONS, 2014f).

4. When presenting annual inflation rates in the following tables and figures, this paper does not include data for 2014 as price data is only available up to October 2014.

5. The 'equivalisation' process adjusts household specific expenditure and income to take account of household composition and is based on the OECD-modified scale equivalisation factors used in the ONS publication on the Effects of Taxes and Benefits (ONS, 2014b). Conceptually, this process accounts for the fact that households with more members are likely to need a higher income to achieve the same standard of living as households with fewer members. However, while a household with two people in it will need more money to sustain the same living standards as one with a single person, the two person household is unlikely to need double the income. It is on this basis that households are divided into deciles for distributional analysis.

6. The respective sub-sections in Section 5 contain the relevant definitions for these sub-groups.

7. In particular, the resulting series are double-chain linked – once in January (to reflect changes in the expenditure weights) and once in February (to reflect changes in the products included in the CPI) (ONS, 2014a).

8. To be precise, we weight household spending from the LCF, and calculate the share of weighted expenditure accounted for by each observed household. This fraction of the CPI expenditure total is allocated to that household. Note that this implicitly means that this dataset adopts the concepts and expenditure definitions of the CPI.

3 . Section 4. Context & literature

A wide range of papers have calculated price indices and inflation rates for population sub-groups in a variety of countries for different time periods. This section first examines the papers that consider the UK evidence, before turning to inflation estimates from a range of other countries.

4.1 UK studies

One of the earliest UK papers to calculate sub-group price indices is Crawford (1994). Looking at the period 1979 to 1992, he found a maximum spread of inflation of 1.6 percentage points between the richest and poorest households. However, while the spread could be high within each year, the differences over the whole period were small. Households in the poorest and richest income groups frequently switched between experiencing the highest and lowest inflation rates and consequently balanced out any within year variation. He concluded that the extent of inflation differences between groups varies depending on the period studied, and on the respective prices of luxuries and essentials and the impact they have on different sub-groups.

The studies that have followed – see, for example; Levell and Oldfield (2011), Pike et al., (2008), Crawford and Smith (2002), Adams et al., (2014) – have all reached similar conclusions. The various periods that have been studied have seen different sub-groups of the population experience different rates of inflation. In particular, Crawford and Smith (2002) found statistically significant differences in average inflation rates experienced by income deciles in most of the 25 years studied, with a maximum spread of around 2 percentage points. However, they conclude that no single group has consistently experienced higher or lower inflation than average. Among these previous studies, analyses that focus on the same period as this work (2003-2014) are particularly relevant. These have concluded that the lower income deciles experienced high inflation rates in years that saw large rises in the price of food and fuel, for example 2006 and 2008 (Levell and Oldfield, 2011).

The Living Cost and Food Survey (LCF) dataset used in this article has been used in various forms in most UK studies looking at sub-group price indices, reflecting the fact that, while there are limitations to this source (see Section 8), it is the most complete dataset that is available to look at spending patterns of UK households. However, the methods used to calculate inflation rates for sub-groups differ from paper to paper. Most previous work uses the observed spending of a given household in the LCF, then 'price updates' that expenditure to the next period – effectively calculating the change in the cost of a fixed consumption basket. This individual household inflation rate is then weighted – either 'plutocratically' or 'democratically' (see Sections 2 and 7) – to produce an aggregate index. However, this method assumes that there is no between-year substitution at the household-level, making it inconsistent with the Consumer Prices Index (CPI), which permits between-year substitution effects (ONS, 2014a). Another method, as presented in Crawford (1994), is to use a Tornqvist index. This form of price index uses expenditure weights from the previous and current periods together to deliver a price index which is more sensitive to substitution effects between years. However, as the LCF is cross-sectional – providing snap-shots of household-level of expenditure in a single period – this approach is less appropriate for our dataset.

Several different sub-groups have been the target of previous analyses. As well as the inflation experience of households in different income deciles discussed above, households with and without children (for example; Crawford, 1994, Crawford and Smith, 2002) and retired and non-retired households (for example; Levell and Oldfield, 2011, Pike et al., 2008, Crawford and Smith, 2002, Leicester et al., 2008) have all been subject to recent work. These studies found that there is no significant difference observed in the inflation experience of households with and without children, with a maximum spread of +/- 0.2 percentage points. Crawford (1994) noted however that the presence of children leads to the household taking on spending patterns similar to that of poorer households: adults forego luxuries and spend more of their budget on goods like food and clothing. By comparison, inflation rates for retired and non-retired households have varied significantly over the last two decades with a maximum spread of just under 3 percentage points. The main drivers behind this spread are rises in food and fuel prices for retired households, and changes in mortgage interest payments for the non-retired households (Leicester et al., 2008).

The conclusions of these papers support the notion that while there are often quite large inflation differentials in specific years, these tend to average out over longer time-frames. Leicester et al. (2008) go further and argue that while inter-sub-group differences are important, intra-group variation is equally substantial. For instance, their analysis – which shows that the various drivers of inflation in recent years has impacted on different types of pensioner households to differing degrees – leads them to the conclusion that it is misleading to talk about a single price index for a sub-group. This thesis is examined more in Section 7.

Finally, a number of previous studies have calculated democratically- and plutocratically-weighted price indices for UK sub-groups and examined the difference between these measures. In particular, Crawford and Smith (2002) found a statistically significant difference between the plutocratic and democratic inflation rates in 18 out of the 25 years studied.

4.2 International studies

The question of different sub-groups of the population facing differing inflation rates was first raised internationally in the late 1950s, when Arrow highlighted the different expenditure patterns of households in the US that lay in different parts of the income distribution, and stated “there should be a separate cost-of-living index number for each income level” (Arrow, 1958). Since then, various studies have been conducted over different time periods and looking at different population sub-groups in various economies. Oosthuizen (2007) presents a summary of international work on sub-group price indices as background for an analysis of inflation experiences in South Africa from 1998 to 2006. This developed earlier work by Ley (2005), which summarised twelve studies examining the inflation rates experienced by different sub-groups from countries such as the US and Argentina. Supporting the findings of recent UK research, both these papers found that no single sub-group consistently experienced a higher or lower inflation rate relative to other groups in the long run. However, within years, there were statistically significant differences experienced by households in different sub-groups.

More recent work includes Hait and Janský (2014) who examined the inflation experiences of households in the Czech Republic in the period 1995 to 2010. They found that only around 60% of households experienced inflation similar to the national average, with higher inflation rates experienced by pensioner households and those with low incomes. This is a theme that runs through much of the literature; households that spend a large proportion of their budget on goods and services that are exposed to large price rises face higher inflation rates. In periods where prices for luxuries are rising at a faster rate than for essentials, this is likely to be high income households, while in periods when the cost of household essentials is rising more quickly, this is likely to be low income and retired households who have limited capacity to substitute towards cheaper products. Much of this work consequently finds that inflation differentials are particularly sensitive to changes in the price of fuel and energy (see, for example, work on US inflation by Hobijn and Lagakos (2005), and work on Austrian inflation by Fritzer and Glatzer (2009)).

In common with this paper, several international studies acknowledge the possibility that differences in sub-group inflation rates could also be caused by price differentials for specific goods across sub-groups. In particular, Oosthuizen (2007) states that “prices are collected from outlets that are generally chosen to be representative of the official population, while this is unlikely to be the case for a specific sub-group”. Ley (2005) outlines how some national statistical offices have looked to overcome this problem. In particular, the Indian Ministry of Statistics and Programme Implementation collect distinctive prices for every sub-group they produce inflation rates for, based on a sample of shops visited by each respective sub-group. Other national statistical offices – see, for example; Australian Bureau of Statistics (2014) and US Bureau of Labor Statistics (2012) – produce sub-group indices based on prices sourced directly from the CPI. However, the Australian Bureau of Statistics make exceptions in cases where it is known that different sub-groups face different prices, such as subsidised public transport fares and pharmaceuticals for retired households.

A number of these international studies also examine the differences between plutocratic and democratic price indices, both for whole economies and for sub-groups within their populations. Oosthuizen (2007) provides a useful summary of the alternative measures, how they differ and in which circumstances each is more appropriate or useful. Most recent papers reference her argument that a plutocratic average is useful for understanding inflationary pressure in the broader economy because the contribution of higher expenditure households is “in line with the overall structure of consumer spending”. By contrast, democratic weights provide a better understanding of the inflation rates faced by different sub-groups, and can be used to capture the ‘average’ household within each sub-group. Most recent studies have therefore gone on to calculate both the plutocratic and democratic measure of inflation and found there are significant differences between them. Ley (2005) concluded the sign and magnitude of this difference varies across country and by year.

Finally, reflecting the variety of the UK literature, a range of different methods has been adopted to calculate inflation rates in studies of other economies. Only one paper (Fritzer and Glatzer, 2009) adopts a comparable approach to that used in this analysis, focussing on inflation rates in Austria. However, after aligning the underlying household expenditure data with the weights for the CPI, their paper excludes the class-level categories which are distorted by this matching technique.

4. 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