Calculating a Retrospective Superlative Consumer Prices Index for the UK

Gareth Clews\(^1\), Ria Sanderson and Jeff Ralph

Summary

The main research aim for this analysis is to calculate a version of the CPI using a superlative index number formula. This is of interest for two reasons; firstly, the test and stochastic approaches to index numbers show that symmetric index formulae have a number of advantages over asymmetric formulae and secondly, a superlative index number formula can approximate a cost-of-living index\(^2\).

It is not possible to calculate a superlative index in a timely manner as this requires expenditure weights from the current time period which are not available. It is, however, possible to calculate a superlative CPI retrospectively once expenditure information for the price collection period becomes available.

The aim of the analysis is therefore to determine what difference is seen, in terms of the index level and the growth in the index, between compiling the CPI using the current practices for locally collected prices and compiling index numbers using superlative formulae for the same data.

Results show that 12 month growth in the locally collected “all items” CPI exceeds the 12 month growth in both the Fisher and Törnqvist superlative price indices. The 12 month growth in the CPI (locally collected data) exceeds the Törnqvist by an average of 0.46 percentage points. Similar results were seen at the COICOP-12 level.

The geometric Laspeyres, Young and Lowe price indices have been proposed as indices which can approximate a superlative index without requiring current period weights. The analysis has therefore examined whether geometric formulae can be used to approximate superlative index numbers in a timely manner.

Results for the locally collected “all items” CPI show that the geometric Young is close to the Törnqvist. The difference between the 12 month growths in the Törnqvist and the geometric Young was between -0.11 and 0.23 percentage points, with an average of 0.06 percentage points. However this relationship is not as clear at the COICOP-12 level where the geometric Young index does not always provide a good approximation of the Törnqvist price index.

\(^1\) gareth.clews@ons.gov.uk
\(^2\) Superlative indices are a subset of symmetric indices, where “symmetric” refers to the treatment of weights in the index formula
1. Introduction

The traditional approach to calculating a consumer price index is to choose a representative basket of goods and services whose content is fixed at the start of each calendar year. Prices of these goods and services are collected every month and price indices calculated at the upper levels using a Lowe Index, which is a weighted average of lower level aggregates, where the weights are the relative expenditure shares of the commodities collected primarily from household expenditure surveys. This approach is familiar as part of the “cost of goods” approach to estimating average price change.

The test approach to Index Numbers suggests that formulae which use weighting information from both the base and current periods in a symmetric way\(^3\) have advantages over asymmetric formulae like the Lowe Index. The economic approach incorporates the effect of consumers adjusting their purchases as relative prices change – this also leads to formulae that use weighting information from both the base and current periods. This behaviour is included within the “cost of living” conceptual framework and is judged by some experts to lead to a preferred measure of price change. The most familiar such formulae are the Fisher and Törnqvist index formulas\(^4\).

---

3 Referred to as Symmetric Indices
4 These are called superlative indices; superlative indices are a subset of symmetric indices
Calculating a Retrospective Superlative Consumer Prices Index for the UK

The Laspeyres index can be shown, under certain conditions, to be an upper bound for a cost of living index. Accordingly, the use of a Laspeyres-type\(^5\) formula has led commentators to state that the traditional approach overestimates inflation, or is "upwardly biased". This was certainly the view taken by the influential Boskin Commission report of 1996, which identified a number of types of measurement error, including not accounting for consumer substitution behaviour.

A number of NSIs have investigated the size of the difference between a Laspeyres-type formula and a Fisher or a Törnqvist formula. In particular, the Bureau of Labour Statistics has been producing a Törnqvist version since 2002 (Cage et al., 2003). The work in this paper investigates this difference for a subset of the UK CPI.

There is a complication in producing Fisher and Törnqvist indices which is that they require weights from the comparison period. These weights are not available in a timely fashion, so these indices can only be calculated in a retrospective form, that is, at least two years after the period to which they refer in the case of the UK. These formulae are therefore not suitable for use by an NSI when producing a monthly set of price index numbers and growths.

This limitation has led to research to see whether it is possible to approximate a superlative index through the use of geometric formulae which do not rely on comparison period weights (for example: de Haan, Balk & Hansen 2009; Lent & Dorfman 2009; Armknecht & Silver 2012). In some cases, these geometric formulae have shown to be a close match to a superlative index. This paper therefore also calculates geometric Laspeyres, geometric Young and geometric Lowe indices and compares them to the superlative indices.

This report details the methodology used to calculate retrospective Fisher and Törnqvist indices and their geometric approximations for a subset of the UK Consumer Prices Index (CPI). Section 2 describes the current CPI methodology, Section 3 details the research questions posed in this analysis, Section 4 describes the applied methodology and results are presented in Section 5. Quality assurance of the analysis is described in Section 6 and Section 7 details further analysis at the COICOP-12 level, which is one level below the all-items CPI. Section 8 tests the sensitivity of the results to the assumptions made during the analysis and conclusions are drawn in Section 9.

2. The UK Consumer Prices Index

The CPI is the UK version of the Harmonised Index of Consumer Prices (HICP); it is compiled in line with European regulations and it is used as the UK Government’s main inflation target. The CPI measures the change in the prices of goods and services bought for consumption by households between a reference period and a comparison period. In this section, a brief summary of the methodology used to calculate the UK CPI is presented, focussing on those areas particularly relevant to this analysis. More detailed information can be found in the CPI Technical Manual (ONS, 2014a).

The CPI is compiled hierarchically; individual price quotes for products are first collated into individual item indices, which in turn are collated into class indices. These class indices are combined to calculate the “all items” CPI. Expenditure information is available and is used to

\(^{5}\) It has been noted by some commentators that it is more appropriate to describe the Laspeyres index as a Lowe-type index; however, the more widely used “Laspeyres-type” description is adopted here.
calculate weights at all but the lowest level of the aggregation structure. In line with the European
regulations for HICP, the UK CPI is calculated using a Lowe formula at those levels where
expenditure information is available.

The items that are priced and used to compile the CPI are known collectively as the “basket of
goods and services” or simply: “the basket”. The basket contains a selection of goods and services
which are representative of consumers’ purchases. Each year the items in the basket are reviewed
and the list of items is updated if necessary (see for example, ONS, 2014b). The updating of the
basket ensures that the items remain representative as expenditure patterns of consumers
change. For example, in 2005, music downloads were added to the basket as a new item to
represent the increasing expenditure on this type of product.

Once price quotes have been collected they are aggregated to the different levels of Classification
of Individual Consumption by Purpose (COICOP). COICOP is an international hierarchical
classification system for categorising consumer expenditure and the aggregation structure of the
CPI follows this classification.

The CPI is compiled from the lowest level upwards, aggregating the appropriate components from
each level to the next. The lowest level is the stratum level; there is no weighting information at this
stage due to the difficulty and cost of collecting such information so price quotes are combined
using unweighted index number formulae. For the most part, this is the geometric mean of price
relatives (Jevons); for the remainder the ratio of average prices (Dutot) formula is used. The
individual strata cover different combinations of region and shop type. Above the stratum level,
weights are applied at each level of the aggregation. Stratum indices are aggregated to the item
level through the application of stratum weights. These weights reflect consumer expenditure by
region and shop type and are updated from a variety of sources which includes the Living Costs
and Food survey (LCF). Indices at the item level are aggregated to the class level through the
application of item weights. Item weights are updated annually from a variety of source data
including the LCF and market research information (ONS, 2014a) with new item weights coming
into the CPI in February. The introduction of these new item weights coincides with the updating of
the representative basket of goods for which the price data are collected. A depiction of the
hierarchical structure is shown in figure 1. For a full description the reader is referred to the

Indices at the class level are aggregated to the all items CPI level through the application of class
weights. As a result of the time required to collect and compile the class weights, these are
introduced into the CPI with a lag from the period from which the data were collected. The class
weights are calculated from National Accounts Household Final Monetary Consumption
Expenditure data from two years prior to the current CPI year. Class weights are price updated to
the current year. This assumes that the quantities bought have not changed from the weight
reference period\(^6\) to the latest price collection period (ILO, 2004\(^7\)). Class weights are re-calculated
annually with the new class weights being introduced into the CPI in January.

\(^6\) The period of time when the expenditure data used to calculate the weights were collected.

\(^7\) Paragraph 1.273.
The CPI is chain linked twice a year to account for the update of the class weights in January and the update of the basket of representative items in February. Chain linking is performed in order to produce an index series that is continuous and on the same scale. The chain linking formula is shown in Annex A; more information can be found in ONS (2014a).

3. Research aims
This analysis sought to answer the following two research questions:

Research question 1 – How does a retrospective superlative CPI compare to the CPI compiled using the current methods?

The main research aim for this analysis was to calculate a version of the CPI using a superlative index number formula. This is of particular interest as the use of a superlative index number formula can represent a cost-of-living index (ILO 2004\(^8\)). The use of superlative, and more generally symmetric index number formulae is also recommended by other approaches to index numbers such as the Test approach.

---

\(^8\) Paragraph 1.97 and references therein.
It is not possible to calculate a superlative index in a timely manner as this requires expenditure weights that refer to the same period as the relevant price data. It is however possible to calculate a superlative CPI retrospectively, once expenditure information for the price collection period becomes available. This is the approach taken by the US Bureau of Labor Statistics (Cage et al 2003).

The aim of the analysis was therefore to determine what difference is seen, in terms of the index level and the growth in the index, between compiling the CPI using the current practices and compiling the CPI using superlative index number formulae.

Research question 2 – Can superlative indices be approximated in a timely fashion, possibly through the use of geometric formulae?

The fact that the Fisher and Törnqvist indices require current period weights means that they can only be produced retrospectively. The geometric Laspeyres has been proposed as an index which can approximate a superlative index without requiring current period weights; it uses the same information as the conventional arithmetic Laspeyres. The analysis has therefore examined whether geometric formulae can be used to approximate superlative index numbers in a timely manner.

4. Methodology

This section describes the methods used to address the research questions described in Section 3. Data availability and processing considerations led to the analysis being performed for only locally collected price quotes; which account for around 57% of the weight of the UK CPI. More details on the data used for the analysis are provided in section 4.1. The analysis required the construction of the CPI using the locally collected price quotes and the current methods (described in Section 2). In addition, addressing the proposed research questions involved the calculation of a variety of indices, which are summarised in table 1.

<table>
<thead>
<tr>
<th>Aggregate formula</th>
<th>Research question</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laspeyres</td>
<td>1</td>
<td>Required in calculation of Fisher index. Can only be calculated retrospectively (requires base period weights).</td>
</tr>
<tr>
<td>Paasche</td>
<td>1</td>
<td>Required in calculation of Fisher index. Can only be calculated retrospectively (requires current period weights).</td>
</tr>
<tr>
<td>Fisher</td>
<td>1</td>
<td>Can only be calculated retrospectively (requires calculation of Laspeyres and Paasche).</td>
</tr>
<tr>
<td>Törnqvist</td>
<td>1</td>
<td>Can only be calculated retrospectively (requires base period and current period weights).</td>
</tr>
<tr>
<td>Geometric</td>
<td>2</td>
<td>Can only be calculated retrospectively</td>
</tr>
</tbody>
</table>

The prices of items in the basket are collected in one of two ways: centrally or locally. A price that is collected locally involves a price collector going to an outlet and recording the price for an item that matches a product description. A price that is centrally collected requires no field work as prices are provided directly to ONS or collected from the internet. The majority of prices used in the CPI calculation are locally collected. The data for locally price quotes was more readily available; in principle the analysis could be extended to cover the whole of the CPI.
Laspeyres (requires base period weights).

Geometric Young 2 Corresponds to geometric formula with price updating removed.

Geometric Lowe 2 Corresponds to using same weights as CPI.

Lowe Both Mimics the CPI methods, used for comparison.

The indices described in Table 1 were all calculated using the Jevons (geometric average of price relatives) formula\(^\text{10}\) at the elementary aggregate level; this mirrors the practice in the CPI. The mathematical formulae for each of the aggregate level indices described in Table 1 are given in Annex B. The following sections describe the methods applied at each level of aggregation to calculate the CPI variants described above.

### 4.1 Price data

The price data used for the calculation of the price indices in this analysis consisted of locally collected price quotes. Initial data cleaning that mirrors the practice in the CPI was performed on these data. Any prices for products deemed to be non-comparable were removed from the analysis. These quotes are identified using markers applied to the data during price collection. Price relatives meeting any of the following criteria were removed from the calculation of the indices:

1. Observations with a zero base price and non-zero current price
2. Observations with a non-zero base price and a zero current price
3. Observations with a zero base price and a zero current price

The above data cleaning removed typically 0.01 percent of the price quotes in any one year. Price relatives greater than ten were also removed. This only affected one of the three years analysed and resulted in the removal of five price quotes.

The price data for January, February to December, and January based on December were treated separately for computational efficiency. There were two motivating factors for taking this approach. Firstly, to replicate the chain linking procedure that takes place in the CPI a separate index is required for each of these periods of time. Secondly, the application of different aggregate level formulae meant that different weights were required for each period of time, and the separate treatment reduced some of the complexity in the analysis. Three years of price data were analysed; this covered the period 2007 to 2009 inclusive. Due to the retrospective nature of the calculation, this required using the weights referring to price collection periods from 2006 to 2011\(^\text{11}\) (inclusive).

---

\(^{10}\) The Jevons is described mathematically as

\[
J_i(t) = \left( \prod_{i=1}^{N} \frac{p_i^t}{p_i^0} \right)^{\frac{1}{N}}
\]

where \(p_i^t\) is the price of item \(i\) in time \(t\), \(p_i^0\) is the price of item \(i\) in time 0 and there are \(i = 1, ..., N\) items. That is, the Jevons is the geometric mean of price relatives.

\(^{11}\) Note that the calculation of the Paasche in year \(y\) requires class weights from \(y\), which in practice are not available until year \(y+2\).
4.2 Weights

The compilation of the CPI relies on three different sets of weights that are applied at the three aggregation levels. These are the:

- stratum weights, which are applied to stratum indices to calculate item indices
- item weights, which are applied to item indices to calculate class indices
- class weights, which are applied to class indices to calculate the all items CPI

In the CPI, the class weights used in year \( y \) are derived from data collected in the calendar year \( y-2 \); this information is then price updated. The item and stratum weights are derived from data collected over the 12 months to June \( y-1 \).

Laspeyres indices and geometric Laspeyres indices rely on weights from the base period, which is January of the current year. Paasche indices rely on weights from the current period. However, only annual weights, rather than monthly weights, are available. This means that for index numbers calculated in this analysis, annual weights are used as a proxy for monthly weights. This is of particular importance for the calculation of those indices that require current period weights (the Paasche, Fisher & Törnqvist). The calculations are therefore not exact versions of these indices. They are the closest calculations possible given the availability of the weights. The sensitivity of the results of the analysis to the use of annual weights as a proxy for monthly weights was investigated and is described in more detail in Section 8.

The weights applied at each aggregation level depend on the aggregation formula being used, and due to the nature of the annual item and class weight updates, the period of time for which the index is being calculated. An attempt was made to match the most relevant weights to each period, given the weight availability and the methods used to construct the CPI. That means that the updating of the item weights in January to allow for the new class weights was a feature of the methods that was maintained throughout the analysis. Similarly, the introduction of new class weights and the introduction of new item weights took place in January and February respectively. The rationale for choosing which weights to apply for each period and each aggregation formula is shown in Figure 2.

In an effort to simplify some of the calculations, a single overall weight was calculated to aggregate from the elementary aggregate indices to the all items CPI. This weight was calculated as the product of the corresponding stratum weights, item weights and class weights. The corresponding mathematical formulae for these combined weights are shown in Annex C. Note that the geometric Laspeyres uses the same weights as the Laspeyres. The geometric Lowe and geometric Young use the same weights as the Lowe and Young indices (the CPI weights and the CPI weights with price updating removed respectively). The Törnqvist uses both a base period and a current period weight. In practice this was achieved by using the same weights as the Laspeyres for the base period weights, and the same weights as the Paasche for the current period weights.

The use of the original class level expenditure data allowed the recalculation of the class level expenditure weights without price updating, meaning that they correspond to the period in which they were collected.
Calculating a Retrospective Superlative Consumer Prices Index for the UK

4.3 Calculation of indices

In line with the current CPI methodology for the locally collected items, shop weights\textsuperscript{12} were applied to the cleaned locally collected price quotes and stratum indices were calculated using the Jevons formula. These were aggregated using the combined weights (described in Section 4.2) and the appropriate aggregate level formulae. Chain linking was applied to each of the indices described in Table 1 at the “all items” level to give a continuous index series from January 2007 to December 2009 inclusive. Currently the CPI has a reference period of 2005; this is where the index series is set to equal 100. In this analysis the first period for which the indices were calculated was January 2007 hence the reference period is January 2007 = 100.

5. Results

This section presents the results of the analysis to address the four research questions described in section 3.

5.1 Retrospective superlative indices for the UK CPI

This section presents the results of calculating superlative indices for the UK CPI, to support research question 1.

5.1.1 Fisher and Törnqvist indices

Retrospective Fisher and Törnqvist indices were calculated for the period from January 2007 to December 2009 inclusive. The index series, the month on month growth and the 12 month growth are shown in Figure 3. These are compared to the Laspeyres and Paasche indices.

\textsuperscript{12} The shop weights are replication factors that are used to reflect the market share of shops with national pricing policies and the market share of retailers with multiple stores in a region (ONS, 2014a).
Figure 3 shows that the Laspeyres increases at the fastest rate, and the Paasche increases at the slowest rate. This is in line with standard expectations where the Laspeyres and Paasche form upper and lower bounds to the Fisher index.

Figure 3 shows that the month on month percentage change is similar for all four indices, apart from some differences at the turning points. The pattern for the 12 month change is very similar between the three indices but the magnitude of the growth differs, with the Laspeyres showing the largest 12 month growth and the Paasche showing the smallest 12 month growth.

Perhaps the most striking result in Figure 3 is the similarity between the Fisher and Törnqvist indices, both in terms of the index series and the growths. This is not unexpected; Diewert (1978) indicated that superlative indices approximate each other to the second order. The difference in 12 month growth between the Fisher and the Törnqvist ranges from -0.009 to 0.017 percentage points (the mean is 0.003 percentage points).

Figure 4 shows the difference between the Fisher and each of the Törnqvist, Laspeyres and Paasche indices in terms of both the month on month growth and the 12 month growth. Note that a positive number means that the growth exceeds the growth in the Fisher index at that point in time. The largest differences between the Laspeyres and the Fisher and the Paasche and the Fisher are observed in January of both 2008 and 2009. The Törnqvist is closest to the Fisher index, both in terms of month on month growth and 12 month growth.

**Figure 3:** Index series, month on month growth and 12 month growth for the Fisher, Törnqvist, Laspeyres and Paasche aggregate level indices
5.1.2 Comparison to the CPI

Figure 5 shows the difference between the Fisher index, the Törnqvist index and the Lowe index (which replicates the current CPI methods but using the locally collected data only). As expected from its use of a Lowe formula at the higher aggregation levels, Figure 5 shows that the CPI exceeds both the Fisher and Törnqvist indices.

The 12 month growth, the CPI exceeds the Fisher by an average of 0.46 percentage points. The largest difference between the CPI and the Fisher index in the 12 month growth is 0.62 percentage points; the smallest difference is 0.32 percentage points. In terms of the month on month growth in the CPI (locally collected data) exceeds that for the Fisher by an average of 0.03 percentage points; the difference in month on month growth between the CPI and the Fisher ranges from -0.05 to 0.15 percentage points.

The 12 month growth, the CPI (locally collected data) exceeds the Törnqvist by an average of 0.46 percentage points. The largest difference observed is 0.60 percentage points and the smallest difference is 0.32 percentage points. In terms of the month on month growth in the CPI (locally collected data)
collected data) exceeds that of the Törnqvist by an average of 0.03 percentage points; the difference in month on month growth ranges from -0.05 to 0.15 percentage points.

**Figure 5: Comparing the Fisher index, the Törnqvist index and the CPI (locally collected data)**

![Graph comparing Fisher, Törnqvist, and CPI indices]

### 5.2 The use of geometric formulae

The difficulty with using superlative index formulae for consumer price indices is the requirement for timely weights, which are not possible to produce in practice. Research question 2 posed the question of whether geometric index formulae could be used as approximations to superlative index numbers. This is a topic that has been explored by other authors (for example, van Kints & Bishop 2013, Armknecht & Silver 2012, Lent & Dorfman 2009).
5.2.1 The geometric Laspeyres as an approximation to a superlative index

Figure 6 compares the index series and shows the difference in month on month growth and 12-month growths from the Fisher index, for both the Törnqvist and the geometric Laspeyres.

**Figure 6:** Comparing the geometric Laspeyres to the Fisher and Törnqvist indices

Figure 6 shows that an index calculated using the geometric Laspeyres formula appears very close to that calculated using both the Fisher and Törnqvist formulae. Note that the indices calculated in this analysis are not a true Fisher, Törnqvist or geometric Laspeyres due to the lack of monthly weights.

The difference in the month on month growth between the Fisher index and the geometric Laspeyres index ranges from -0.03 to 0.02 percentage points over the period considered, with an average of 0.00008 percentage points. In terms of the twelve month growth, the difference ranges
Calculating a Retrospective Superlative Consumer Prices Index for the UK

from -0.04 to 0.05 percentage points, with an average of 0.005 percentage points. Previous work carried out by ONS on the RPI found a similar result (Wood, 2001).

The difference in the month on month growth between the Törnqvist and the geometric Laspeyres ranges from -0.02 to 0.02 percentage points over the period considered, with an average of -0.0002 percentage points. In terms of twelve month growth, the difference ranges from -0.03 to 0.03 percentage points, with an average of 0.002 percentage points. The sensitivity of these results to monthly variations in the weights has been examined and is described in section 8.

5.2.2 The use of geometric formulae that can be calculated in a timely manner

The geometric Laspeyres index shown in Figure 6 used expenditure weights from the base period, meaning that like the Fisher and Törnqvist, this index can only be calculated retrospectively. Geometric Lowe and geometric Young formulae can however be calculated in a timely manner. The first relies on the same weights as the CPI, and the second relies on these same weights but with price updating removed.

Figure 7: Comparing the geometric Lowe and geometric Young to the CPI (locally collected data only) and the Törnqvist index
Whereas an arithmetic Lowe may be preferred to an arithmetic Young (ILO, 2004\textsuperscript{13}), the geometric formulation of the Lowe is not a fixed basket index and has no support in the literature (see Armknecht & Silver 2012). In contrast, Armknecht & Silver (2012) identify that a geometric Young index is an appropriate index.

Figure 7 compares the CPI (locally collected data) and the Törnqvist index to a geometric Lowe index and to a geometric Young index. The Törnqvist has been chosen over the Fisher here for two reasons. Firstly, the Fisher and Törnqvist appear to approximate each other very well, and secondly, the Törnqvist is the geometric mean of the geometric Laspeyres and geometric Paasche, and as such has a comparable algebraic form to the geometric indices under consideration.

The effect of using the price updated weights in the geometric formula (geometric Lowe) is to move the index consistently above the Törnqvist index; however this difference is small. In terms of month on month growth, the difference between the Törnqvist and the geometric Lowe ranges from -0.13 to 0.09 percentage points, with an average of -0.006 percentage points. In terms of 12 months growth, the difference between the Törnqvist and the geometric Lowe ranges from -0.16 to 0.06 percentage points, with an average of -0.04 percentage points. This indicates that in terms of the growths, the geometric Lowe typically exceeds the Törnqvist; however this difference is generally small.

The geometric Young also appears to very closely follow the Törnqvist. In terms of month on month growth, the difference between the Törnqvist and the geometric Young ranges from -0.08 to 0.11 percentage points, with an average of 0.002 percentage points. In terms of 12 month growths, the difference between the Törnqvist and the geometric Young ranges from -0.11 to 0.23 percentage points, with an average of 0.06 percentage points. Again, the difference is small and is much smaller than the difference between the CPI and the Törnqvist index.

In comparison, the difference in the 12-month growth between the Törnqvist and the CPI (locally collected data) and ranges from -0.60 to -0.32 percentage points, with an average of -0.46 percentage points. It seems from Figure 7 that the use of a geometric Lowe formula or a geometric Young formula could lead to an index that is a close approximation to a Törnqvist index in terms of the 12 month growth. Note however that the geometric Lowe has no conceptual support (see Armknecht & Silver 2012) which might lead to the geometric Young being preferred.

6. Quality assurance of methods

The methods applied in this analysis have been quality assured by comparing the calculated indices to the published CPI and by comparing the results with similar research in the literature.

6.1 Comparison to published CPI growths

One key part of the quality assurance process was to check whether the growths in the CPI calculated here from locally collected data are consistent with the published CPI growths. Figure 8 compares the published 12 month growths for the CPI to the 12 month growths calculated from the

\textsuperscript{13} Paragraph 1.37 and references therein.
Calculating a Retrospective Superlative Consumer Prices Index for the UK

locally collected CPI data in this analysis\textsuperscript{14}. The two do not match; this is not unexpected as the CPI calculated here excludes the centrally collected items, which include some services and products such as energy prices. However, the general shape of the two series is similar, which provides some assurance of the methods in place for this analysis. Further analysis is required to ensure that the differences seen are due only to the exclusion of the centrally collected items from this analysis, but this is beyond the scope of the analysis presented here.

\textbf{Figure 8: Comparing 12-month growths for the CPI (locally collected data) to the published CPI}

6.2 International comparisons

A number of National Statistics Institutes (NSIs) have published retrospective superlative consumer price indices. Most notably, the US Bureau of Labour Statistics has been publishing a retrospective Törnqvist version of the consumer prices index (the “chained C-CPI-U”) for more than ten years. Over that period, the yearly percentage difference between the published CPI (the “CPI-

U") and the retrospective Törnqvist index was typically between 0.2 and 0.3 percentage points, with a maximum of 0.45 percentage points in 2005 and a minimum of -0.13 percentage points in 2008 (Shoemaker, 2013). These differences are typically smaller than the equivalent comparison in this analysis, which found that the average difference between the CPI (locally collected data) and the Törnqvist was approximately 0.46 percentage points, with a range from 0.32 to 0.60 percentage points over the period of interest. There are however differences in methodology that may explain this difference in the results. The US chained C-CPI-U combines lower level indices calculated using geometric Laspeyres and Laspeyres formulae into a higher level Törnqvist formula (Cage, Greenlees & Jackman, 2003). In contrast, the Törnqvist in this analysis was applied at a much lower level, to combine elementary aggregates using a combined weight to aggregate from the elementary aggregate to the all items CPI.

Superlative consumer price indices have also been calculated by Statistics New Zealand (Statistics New Zealand, 2011) and by van Kints & Bishop (2013) using Australian data. Statistics New Zealand (2011) finds the typical difference between inflation rates calculated using the Laspeyres and Fisher to be 0.1 percentage points. Again, this is smaller than the size of the effect observed in this analysis. Similarly, Van Kints & Bishop (2013) find an average difference between a Laspeyres and a Fisher over the period 2000 to 2011 of 0.24 percentage points. These differences are both typically smaller than the difference seen in this analysis. This could be due to the frequency of weight updating in each analysis. Statistics New Zealand (2011) describe how the CPI is reweighted in general every three years and van Kints & Bishop (2013) describe how expenditure data is collected every six years. In contrast, the weights used in this work are updated annually so have the potential to better reflect the current period weights required by the superlative formulae.

In addition, it is important to note that this analysis does not cover all commodities in the full UK CPI; for computational reasons the analysis has been restricted to the locally collected price quotes only. Although Section 6.1 indicates that the CPI (calculated using locally collected data) shows similar 12-month growths to the UK published CPI, there are some differences. Further work would be required to ascertain whether the relationships identified in this research persist for the full CPI sample of price quotes. It would also be possible to test whether the differences in methodology used by other NSIs can explain the differences in the results. This would involved testing whether aggregating lower level Laspeyres indices with a superlative index formulae and whether less frequent updating of the weights could lead to a smaller difference between the CPI (locally collected data) and the Törnqvist. This would require substantial further analysis.

7. Analysis at the COICOP-12 level

One of the main results from Section 5 is the apparent success of the geometric formulae (geometric Laspeyres, Lowe and Young indices) at approximating the Törnqvist formula. If this is indeed the case, the implication is that index number formulae that can be calculated in real time could be applied in consumer price indices to approximate a symmetric price index and therefore also a cost of living index.

There are however two conditions surrounding this result. Firstly, the analysis is based only on the highest level of aggregation in the CPI. Secondly, the analysis was carried out over a short time period where inflation is relatively stable. It is important to therefore assess whether the conclusions change at (a) lower aggregation levels and in (b) different inflationary eras.
These areas can both be addressed by carrying out the same analysis at the COICOP-12 level (one level below the all items CPI). Performing the analysis at this level provides a useful check of the results seen at the overall level and it also offers an opportunity to compare formulae under different economic conditions, as the diversity of commodities at the COICOP-12 level mean that even over this short time period, the indices behave differently.

7.1 Methodology

The methodology applied in this analysis exactly mirrors the original analysis. The key differences with this approach are that rather than calculating a combined weight for each stratum, which indicates the contribution of each stratum to the all items CPI level, these weights were re-calculated so that they represented only the contribution to the relevant COICOP-12 level for each individual stratum. The indices were then calculated in the same way as the original analysis. The assumptions made in the original analysis apply also to this work and this analysis is also based only on the locally collected price quotes.

Table 2: Description of the categories within COICOP-12

<table>
<thead>
<tr>
<th>COICOP-12</th>
<th>Description</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Food and non-alcoholic beverages</td>
<td>103</td>
</tr>
<tr>
<td>2</td>
<td>Alcohol and tobacco</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>Clothing and footwear</td>
<td>62</td>
</tr>
<tr>
<td>4</td>
<td>Housing and household services</td>
<td>115</td>
</tr>
<tr>
<td>5</td>
<td>Furniture and household goods</td>
<td>68</td>
</tr>
<tr>
<td>6</td>
<td>Health</td>
<td>24</td>
</tr>
<tr>
<td>7</td>
<td>Transport</td>
<td>152</td>
</tr>
<tr>
<td>8</td>
<td>Communication</td>
<td>24</td>
</tr>
<tr>
<td>9</td>
<td>Recreation and culture</td>
<td>153</td>
</tr>
<tr>
<td>10</td>
<td>Education*</td>
<td>18</td>
</tr>
<tr>
<td>11</td>
<td>Restaurants and hotels</td>
<td>138</td>
</tr>
<tr>
<td>12</td>
<td>Miscellaneous goods and services</td>
<td>100</td>
</tr>
</tbody>
</table>

1. Weight in all items CPI (parts per thousand in 2007; the first year in the analysis)
2. Excluded from the analysis as no locally collected price quotes

The indices were calculated for 11 of the 12 COICOP domains. The COICOP structure is shown in Table 2. It was not possible to calculate indices for COICOP 10 “Education”, as there were no relevant price quotes in the local collection. The weights of each domain in the “all items” CPI in 2007 is also shown to give an indication of the relevant significance of each domain to the all items CPI; this year was chosen as it refers to the first year in the analysis.

7.2 Results

7.2.1 COICOP-12 indices

Figure 9 shows the COICOP-12 indices calculated using the locally collected price quotes for the period 2007 to 2009 inclusive. This shows that there is differing behaviour in the index series over this time period; for example “Clothing and footwear” shows a decrease in the index over time, and “Communication” shows a decrease followed by an increase. The remaining indices generally increase over the period in question. The largest increase over the period is shown by the “Food and non-alcoholic beverages” category.
7.2.2 Geometric approximations to the Törnqvist

The geometric Laspeyres, geometric Lowe and geometric Young were compared to the Törnqvist index in terms of the 12-month growth, which is the headline rate reported in the CPI statistical bulletin. This comparison is shown in Figure 10, where the difference in the 12-month growth between these geometric indices and the Törnqvist index is displayed. The CPI (calculated using locally collected data) is also shown. In this figure, a difference that is consistently close to zero over time could indicate that the index in question is a good approximation to the Törnqvist. Each panel in Figure 10 shows the 12-month growths for different COICOP-12 categories. These are ordered from bottom left to top right in order of increasing weight within the CPI in 2007. This ordering stays the same from 2007 to 2008 inclusive, with only Transport overtaking Recreation & Culture as the highest weighted COICOP-12 in 2009.

Only the geometric Young and the geometric Lowe could be calculated in real time; the geometric Laspeyres requires up-to-date expenditure weights which cannot be obtained with current practices. This effectively rules out the geometric Laspeyres as a timely approximation of a symmetric index unless more timely weights could be obtained. Therefore, the remainder of the

---

comparisons made in this analysis consider only the geometric Lowe and the geometric Young. Note that the geometric Lowe has no conceptual support (for example, Armknecht & Silver 2012).

In terms of the all items CPI level, the original analysis found that the average difference in the 12 month growth between the Törnqvist and the geometric Lowe was -0.03 percentage points (the range was -0.16 to 0.07 percentage points). The average difference in the 12 month growth between the Törnqvist and the geometric Young was 0.07 percentage points (the range was -0.11 to 0.23 percentage points). In contrast, the average difference in the 12 month growth between the Törnqvist and the CPI (calculated using locally collected data) was -0.48 percentage points (the range was -0.61 to -0.36 percentage points). A similar comparison was made for the COICOP-12 domains and is shown in table 3 alongside the standard deviation of the differences in the 12-month growth between the Törnqvist and each of the geometric Young, geometric Lowe and the CPI (calculated using locally collected data).

**Table 3: Average difference in 12 month growth from Törnqvist in percentage points (standard deviations in brackets)**

<table>
<thead>
<tr>
<th>Category</th>
<th>Geometric Young</th>
<th>Geometric Lowe</th>
<th>CPI (locally collected data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and non-alcoholic beverages</td>
<td>0.22 (0.14)</td>
<td>0.39 (0.15)</td>
<td>-0.42 (0.15)</td>
</tr>
<tr>
<td>Alcohol and tobacco</td>
<td>-0.03 (0.04)</td>
<td>-0.08 (0.15)</td>
<td>-0.25 (0.16)</td>
</tr>
<tr>
<td>Clothing and footwear</td>
<td>0.03 (0.19)</td>
<td>0.01 (0.16)</td>
<td>-0.41 (0.19)</td>
</tr>
<tr>
<td>Housing and household services</td>
<td>0.02 (0.13)</td>
<td>-0.19 (0.07)</td>
<td>-0.37 (0.13)</td>
</tr>
<tr>
<td>Furniture and household goods</td>
<td>0.10 (0.12)</td>
<td>-0.10 (0.10)</td>
<td>-0.59 (0.12)</td>
</tr>
<tr>
<td>Health</td>
<td>0.05 (0.06)</td>
<td>0.02 (0.07)</td>
<td>-0.07 (0.06)</td>
</tr>
<tr>
<td>Transport</td>
<td>-0.16 (0.08)</td>
<td>0.01 (0.10)</td>
<td>-0.17 (0.08)</td>
</tr>
<tr>
<td>Communication</td>
<td>0.11 (0.22)</td>
<td>0.11 (0.16)</td>
<td>-0.02 (0.22)</td>
</tr>
<tr>
<td>Recreation and culture</td>
<td>0.09 (0.20)</td>
<td>-0.07 (0.17)</td>
<td>-0.57 (0.20)</td>
</tr>
<tr>
<td>Restaurants and hotels</td>
<td>0.01 (0.08)</td>
<td>-0.02 (0.06)</td>
<td>-0.09 (0.08)</td>
</tr>
<tr>
<td>Miscellaneous goods and services</td>
<td>0.05 (0.05)</td>
<td>0.01 (0.06)</td>
<td>-0.15 (0.05)</td>
</tr>
</tbody>
</table>

These results indicate that on average, the CPI (locally collected data) exceeds the Törnqvist for each of the COICOP-12 categories investigated in terms of the 12 month growth. The picture that emerged from the original analysis is however not as clear when considering these lower level indices. That is, the geometric Lowe and the geometric Young do not always closely follow the Törnqvist at these levels. This is particularly apparent in the “Food and non-alcoholic beverages” category where both the geometric Lowe and the geometric Young over-estimate the 12-month growth as described by the Törnqvist index.
### 7.2.3 The assumption of a constant basket

The analysis assumes a constant basket over time to ensure that differences in the calculated indices are due only to differences in formulae and not the composition of the basket. To assess the impact of this assumption, a version of the analysis was performed where the basket was...
allowed to vary on an annual basis. Table 4 shows the average difference between the Törnqvist and each of the geometric Lowe, geometric Young and the CPI (locally collected data) for the case where the basket is allowed to change between years. Table 4 shows that the effect of allowing the basket to change differs between each COICOP-12 category, in some cases increasing the difference between the Törnqvist and the specified formula and in other cases reducing the difference. The main effect appears to be a general increase in the standard deviation of the differences. This indicates that in a situation more like that which would appear in practice, the variation in the difference between the Törnqvist and the other formulae may be larger than the examples shown in Figure 10.

Table 4:  Average difference in 12 month growth from Törnqvist in percentage points (standard deviations in brackets) for a changing basket

<table>
<thead>
<tr>
<th>Category</th>
<th>Geometric Young</th>
<th>Geometric Lowe</th>
<th>CPI (locally collected data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and non-alcoholic beverages</td>
<td>0.12 (0.23)</td>
<td>0.27 (0.40)</td>
<td>-0.55 (0.35)</td>
</tr>
<tr>
<td>Alcohol and tobacco</td>
<td>0.06 (0.12)</td>
<td>-0.12 (0.26)</td>
<td>-0.28 (0.28)</td>
</tr>
<tr>
<td>Clothing and footwear</td>
<td>0.10 (0.16)</td>
<td>0.08 (0.16)</td>
<td>-0.34 (0.18)</td>
</tr>
<tr>
<td>Housing and household services</td>
<td>0.01 (0.20)</td>
<td>-0.20 (0.08)</td>
<td>-0.37 (0.13)</td>
</tr>
<tr>
<td>Furniture and household goods</td>
<td>0.12 (0.14)</td>
<td>-0.08 (0.19)</td>
<td>-0.58 (0.19)</td>
</tr>
<tr>
<td>Health</td>
<td>-0.04 (0.08)</td>
<td>-0.08 (0.09)</td>
<td>-0.16 (0.10)</td>
</tr>
<tr>
<td>Transport</td>
<td>-0.19 (0.07)</td>
<td>-0.00 (0.11)</td>
<td>-0.19 (0.09)</td>
</tr>
<tr>
<td>Communication</td>
<td>0.11 (0.16)</td>
<td>0.11 (0.16)</td>
<td>-0.02 (0.22)</td>
</tr>
<tr>
<td>Recreation and culture</td>
<td>0.26 (0.46)</td>
<td>-0.04 (0.30)</td>
<td>-0.55 (0.29)</td>
</tr>
<tr>
<td>Restaurants and hotels</td>
<td>0.00 (0.05)</td>
<td>-0.03 (0.05)</td>
<td>-0.10 (0.07)</td>
</tr>
<tr>
<td>Miscellaneous goods and services</td>
<td>0.23 (0.28)</td>
<td>0.18 (0.28)</td>
<td>0.02 (0.24)</td>
</tr>
</tbody>
</table>

7.3 Interpretation

The results in Figure 10 are mixed; in some COICOP-12 domains, the geometric Young does provide a reasonable approximation to the Törnqvist whereas in others this not the case. This indicates that the main conclusion seen at the all-items CPI level (that a geometric Laspeyres or a geometric Young might provide a good approximation to a Törnqvist) does not always hold at this lower level. Comparing the geometric formulae to the Törnqvist shows that in the case where the expenditure shares applied to the geometric formulae equal the arithmetic mean of the expenditure shares from the base period and the current period for all items, then the formulae will be equal. There are however other solutions that would lead to the formulae being equal as an increase in the weight for one item can be offset by a decrease in the weight for another item. The weight and the price relatives cannot be separated, so as long as the increase in the weighted contribution from one item matches the decreased weighted contribution from another item, the indices could still be equal.

The difference between the Törnqvist and the geometric Laspeyres or geometric Young can be examined in more detail by evaluating the ratio of the formulae. This is an approach used by Armknecht & Silver (2012) to compare the geometric Lowe to the Törnqvist. Balk (2008) also uses a similar approach to compare the geometric Paasche to a geometric Laspeyres, which allows this relationship to be written in terms of the covariance between the changes in expenditure shares and the relative price changes. This is a useful approach as given the available data it can also be explored empirically. Applying this approach to a comparison between the Törnqvist and the geometric Young yields the following expression:
Calculating a Retrospective Superlative Consumer Prices Index for the UK

\[
\ln \left( \frac{P_{t}^{b,j}}{P_{t}^{d,j}} \right) = \sum_{i=1}^{N} s_{i}^{b} \left\{ \frac{1}{2} \left( s_{i}^{0} + s_{i}^{f} \right) \frac{1}{s_{i}^{b}} - 1 \right\} \ln \frac{R_{i}^{b,j}}{P_{GY}^{0,j}}
\]

where \( s_{i}^{0} \) is the base period expenditure share for item \( i \), \( s_{i}^{f} \) is the current period expenditure share for item \( i \), \( s_{i}^{b} \) is the expenditure share in period \( b \) (which precedes the base period) for item \( i \), \( R_{i}^{b,j} = \frac{P_{i}^{b,j}}{P_{i}^{0,j}} \) is the price relative for item \( i \) and there are \( i = 1, \ldots, N \) items. The full derivation appears in Annex D.

In the comparison of the geometric Paasche and geometric Laspeyres formulae explained by Balk, the equivalent expression is described as a covariance between expenditure share changes and relative price changes such that a positive correlation between these, results in the geometric Paasche exceeding the geometric Laspeyres. Applying the same logic to the equation above indicates that the correlation between the term in brackets:

\[
\left\{ \frac{1}{2} \left( s_{i}^{0} + s_{i}^{f} \right) \frac{1}{s_{i}^{b}} - 1 \right\}
\]

and the relative prices:

\[
\ln \frac{R_{i}^{b,j}}{P_{GY}^{0,j}}
\]

could be used to describe how the Törnqvist compares to the geometric Young. Again, following Balk (2008), a positive correlation between these terms would lead to the Törnqvist exceeding the geometric Young and a negative correlation would lead to the reverse. This has been tested empirically by calculating the Pearson’s product moment correlation coefficient between the two terms highlighted above.

Figure 11 compares these correlation coefficients to the logarithm of the ratio of the Törnqvist to the geometric Young. This clearly shows that a positive correlation leads to the Törnqvist exceeding the geometric Young and a negative correlation leads to the geometric Young exceeding the Törnqvist. It also begins to highlight why there are differences between different COICOP-12 categories. The panels are ordered from bottom left to top right in order of the increasing absolute average difference in 12-month growth between the Törnqvist and the geometric Young, such that Restaurants and hotels shows the smallest absolute difference and Food and non-alcoholic beverages shows the largest absolute difference. The differences in the correlations seen reflect how well the geometric Young approximates the Törnqvist.
Figure 11: The logarithm of the ratio of the Törnqvist to the geometric Young by COICOP-12\textsuperscript{16} as a function of the correlation coefficient between the relative change in expenditure shares and the relative change in prices.

\textsuperscript{16} The all items CPI is also included and the panels are ordered from bottom left to top right by increasing absolute average difference in 12 month growth between the Törnqvist and the geometric Young.
8. Sensitivity Analysis

In an effort to establish whether the observed results are a consequence of the methodological decisions made during this analysis, the following key assumptions have been tested by performing a sensitivity analysis:

- The assumption of a constant basket of goods over the period of the analysis (2007 to 2009 inclusive)
- The rescaling of class weights to ensure that the sum of the weights over all classes equals 1000
- The use of annual weights in formulae that require weights for the current period

The following sections address each of these assumptions in turn.

8.1 The assumption of a “constant” basket

The CPI basket of goods and services is updated at the start of each calendar year; this process removes products from the basket that are deemed to no longer represent consumer expenditure and introduces new products that have a growing expenditure share.

The analysis to construct a true Laspeyres price index considered three years of price quotes data (2007 to 2009 inclusive) and due to the different index number formulae used, this analysis required CPI weighting information from a longer time period (2006 to 2011 inclusive). As such, the basket changes over the period of the analysis. To ensure that the only difference being measured in the analysis was that difference due to the use of different index number formulae, only those products in each year that had a weight available for all periods from t-1 to t+2 (inclusive) were included. This amounts to using a constant basket of goods over the period of the analysis.

The negative effect of this methodological choice is to exclude all new products that were brought into the basket over this period. In an effort to understand whether this has an impact on the observed price indices, additional analysis has been undertaken to evaluate how many items have been excluded from the analysis and also to consider the impact on the resulting indices if these items had been included.

8.1.1 Items rejected from the analysis

For each year considered in the calculation of indices, the overlap between the weights and the price quotes were examined. As indicated above, price quotes were rejected if any one of the weights from the period t-1 to t+2 inclusive were missing. If all price quotes for a specified item were excluded, the item was deemed to be “rejected”. Items could also be “partially rejected” if only some of the price quotes were lost (this can occur for specific strata or in a specific month). The valid quotes from partially rejected items were still included in the original analysis.

The number of rejected, partially rejected and the number of included items for each analysis year are shown in Table 5. This indicates that typically 13 to 15 percent of the items in each analysis year are rejected.
Table 5: The number of included items, rejected items and partially rejected items for each analysis year

<table>
<thead>
<tr>
<th>Analysis year</th>
<th>Number of included items</th>
<th>Number of rejected items</th>
<th>Number of partially rejected items</th>
<th>Total number of items</th>
<th>Percentage rejected or partially rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>480</td>
<td>77</td>
<td>7</td>
<td>564</td>
<td>14.8%</td>
</tr>
<tr>
<td>2008</td>
<td>487</td>
<td>67</td>
<td>6</td>
<td>560</td>
<td>13.0%</td>
</tr>
<tr>
<td>2009</td>
<td>486</td>
<td>65</td>
<td>10</td>
<td>561</td>
<td>13.4%</td>
</tr>
</tbody>
</table>

The items highlighted as “rejected” or “partially rejected” in Table 4 were explored in more detail in an effort to understand why these items fell into these categories. For each rejected quote the pattern of missing weights was determined.

Figure 12 shows the number of items with each pattern of missing weights for the three analysis years. For example a weight missing in both t+1 and t+2 would appear in figure 12 as “t+1 & t+2”. The patterns are ordered in Figure 12 from top to bottom by the mean number of items with that missing weight pattern across the three analysis years. Figure 12 shows that the largest average number of rejected items occurs when the t-1 weight is missing. This would be the case for new items as they would not have a weight in the previous year.

Figure 12: The temporal pattern of missing weights for rejected or partially rejected items

8.1.2 The impact on indices of a constant basket

The final part of the analysis of the constant basket involved the calculation of the indices using all price quotes (i.e. including those items that had previously been excluded). This version of the
analysis only rejected price quotes if the weights in all time periods were missing. Due to the different weights used in the applied index number formulae, the result is that the same basket is not used for all formulae. This means that any difference in the formulae now also reflects these differences in the basket.

The resulting indices were compared to the original analysis; this comparison is shown in Figure 13. The majority of the index number formulae are not affected by the use of the “fuller” basket. This implies that the use of the common basket in the original analysis makes little difference to the observed index series. The most noticeable difference is in the series that use the geometric Laspeyres and the Laspeyres, where using the reduced basket slightly lowers the index series in comparison to the full basket.

**Figure 13:** Comparing the “full” basket with the “reduced” (constant) version used for the original analysis

8.2 The rescaling of weights
The analysis made use of only the locally collected price quotes, primarily for computational efficiency. A consequence of this is that the weights do not sum to 1000 (or 100) at each aggregation level. The original analysis therefore included a step to re-scale the stratum, item and class weights to ensure that they do sum to 1000. This effectively divides the "spare weight" proportionally amongst all strata, items or classes. This practice of rescaling the weights was investigated in an effort to see whether this was affecting the results. The analysis was performed with the rescaling and without and the resulting indices were compared.
Figure 14: Comparison of indices calculated with rescaled weights ("yes") and those without rescaling ("no")

Figure 15: Comparing index number formulae without and with rescaling of weights
Figure 14 shows indices calculated with and without this rescaling practice (note that a Jevons elementary aggregate is used throughout). The effect of rescaling the weights is to bring the index series down relative to the version without rescaling; this is true for all index number formulae shown in Figure 14\(^{17}\).

Figure 15 compares the different index number formulae directly for the case where the weights are rescaled (the original analysis) and in the case where the weights are not rescaled. Although Figure 14 shows some differences in the indices with and without rescaling, this seems to affect all formulae to a similar extent. The comparison in Figure 15 also indicates that the conclusions of the original work are unaffected by the removal of the weight rescaling. Note that the Fisher, geometric Laspeyres and Törnqvist lie very close to each other.

### 8.3 The use of annual weights

The CPI weights are derived primarily from National Accounts Household Final Monetary Consumption Expenditure and the Living Costs and Food Survey. These weights are updated on an annual basis, with class weights being updated in January and lower level weights being updated in February. The weights are compiled on an annual basis using these data sources, which means that each weight should be considered as an annual figure. However, index number formulae such as the Paasche, Fisher and Törnqvist rely on having current period weights, which as the CPI is compiled on a monthly basis would mean that monthly weights would be required.

The analysis presented here uses a fixed weight for the current period, meaning that it is only updated once a year. To test whether the conclusions are affected by the use of this fixed current period weight, the use of monthly weights was explored. It may be possible to compile monthly weights from the expenditure micro data, however, this is a time consuming activity. An alternative approach was devised to test the susceptibility of the indices to monthly weights. This relied on perturbing the annual weight by varying amounts to create changing “monthly” weights. The amount of perturbation was varied to assess whether the volatility of the monthly weights also had an impact on the indices.

#### 8.3.1 Methodology for creating monthly weights

The weights in the CPI consist of three components: a stratum weight, an item weight and a class weight. The weights are proportions of expenditure, so at each level of aggregation these proportions are constrained to be one\(^{18}\). For the purposes of calculating superlative indices, the three weights were combined into a single composite weight. Therefore, perturbing any one component of this weight would in effect perturb the overall applied weight. Operationally, it is more convenient to perturb only one component of the composite weight as otherwise it is necessary to constrain the weights at each level. Perturbing the weights at only one aggregation level means that constraining the weights only need to occur once. The perturbation was applied at the class level.

The annual weight for a particular class was used as the basis for calculating “monthly” weights. The annual weight was first assigned to each month of the year. Random noise was then added to

---

\(^{17}\) Note that the removal of rescaling also means that the summed weights can differ from year to year. The summed weights are explicitly included in the index number calculations.

\(^{18}\) In practice the weights are presented as parts per hundred or parts per thousand.
Calculating a Retrospective Superlative Consumer Prices Index for the UK

the annual weight independently in each month to simulate variation on a monthly frequency. The random noise was drawn from a normal distribution. This choice of distribution was motivated by considering the difference between the class weights from one year to the next. This was found to be approximately normally distributed. Fitting a normal distribution to the annual difference in the class weights gave a mean of -0.006 and a standard deviation of 1.5. The starting point for the analysis was therefore to draw numbers at random from a normal distribution with mean zero and a standard deviation of 1.5.

The annual variation in class weights might be considered a rather conservative estimate of monthly variation, as the variation has effectively been smoothed out over the year. To investigate the impact on the indices of class weights that were more volatile, further experiments were carried out drawing noise at random from a normal distribution with a standard deviation of 3 and with a standard deviation of 4.5. The mean remained at zero throughout. The distribution of the annual difference in class weights over the period 2007 to 2009 inclusive is shown in Figure 16, alongside the normal distributions used to simulate the monthly weights.

**Figure 16:** The distribution of the annual difference in class weights, overlaid with the normal distributions used to simulate noise in the monthly weights

Once the noise had been added to the class weights in each month, there were two different constraining procedures that were carried out. The first ensured that the mean of the monthly weights within a class was exactly equal to the annual weight. As noise was drawn from a normal distribution of mean zero and some of the class weights are already small, it was possible to create negative monthly weights. When this occurred, the negative weights were reset to the lowest
possible value (this was set equal to one part per thousand, as no zero class weights were observed in practice). The rescaling of the mean was then carried out again. This occurred iteratively until no negative weights remained. The second procedure constrained the sum over all class weights to equal 1000 in each month. This latter stage was equivalent to the approach that was taken in the original analysis.

These simulated monthly weights were then used in the calculation of those index formulae that require current period weights. As the noise was being drawn at random in each month, it made sense to repeat this procedure to allow different levels of variation to be applied to each class weight. The whole procedure was repeated fifty times to give an indication of how different balances of simulated monthly weights could affect the indices.

8.3.2 Results of applying simulated monthly weights

Figure 17: The median (line) and interquartile range (shaded region) of the differences in 12 month growths between the Törnqvist and the geometric Laspeyres calculated using simulated monthly weights from a normal distribution with differing standard deviations. The difference in 12 month growths between the Törnqvist and the geometric Laspeyres from the original analysis is shown as a dashed black line.
Note that the presence of monthly weights does not only affect those indices that require current period weights. They also impact on the base period weight used for formulae such as the geometric Laspeyres, as this can now refer to a single month rather than the previous annual period. One of the main questions concerning the use of annual weights was whether one of the conclusions of the original analysis, namely the close approximation of the geometric Laspeyres and the Törnqvist, would still hold if monthly weights were available. The simulated monthly weights calculated were used in an effort to answer this question by calculating geometric Laspeyres and Törnqvist indices for each of the 50 iterations of the simulated monthly weights. The difference in 12 month growth rates between the geometric Laspeyres and the Törnqvist were calculated for each iteration. The interquartile range of this difference and the median difference are shown in figure 17 as shaded regions and lines respectively for each of the different distributions used to generate the monthly weights.

Figure 17 shows that as the standard deviation of the distribution used to generate the weights increases, so too do the spread of the differences in the 12 month growths between the geometric Laspeyres and the Törnqvist and also the median difference. The distribution that approximates the annual differences in class weights (standard deviation = 1.5) results in the smallest variation in differences in 12 month growths. However, in all cases, the difference in 12 month growth between the geometric Laspeyres and the Törnqvist is much less than the difference observed in 12 month growths between the CPI (locally collected data) and the Törnqvist in the original analysis. This difference varied between a minimum of 0.32 percentage points and a maximum of 0.60 percentage points over the time period evaluated. These levels are not reached in figure 17.

The analysis indicates that variation in the weights on a monthly basis does lead to larger differences between the geometric Laspeyres and the Törnqvist in terms of 12 month growths than the original analysis, which used annual weights. As the level of volatility in these simulated monthly weights increases, the difference also increases; however, within the range of simulations carried out here this does not reach the level of the difference between the CPI (locally collected data) and the Törnqvist.

This indicates that the similarity seen between the geometric Laspeyres and the Törnqvist may in part be due to using annual weights. This is unsurprising, given that these two formulae will yield equivalent indices if the expenditure shares for items in the base period equal the expenditure shares for items in the current period. The limitation of the approach taken here in simulating monthly weights is that the true monthly variation in weights is unknown. Weights have been simulated based on drawing noise from a normal distribution with an increasing standard deviation, such that the monthly weights are assumed to vary more than the year to year changes. However, monthly expenditure shares would need to be calculated to test whether these are effective assumptions.

9. Conclusions

This paper presents a number of results related to the use of alternative index number formulae at the higher aggregation levels in the Consumer Prices Index (CPI). The results presented in this paper are based only on analysing the locally collected price quotes. The centrally collected prices have been excluded.
Calculating a Retrospective Superlative Consumer Prices Index for the UK

The analysis finds that the CPI (calculated using locally collected data) exceeds both a Fisher and a Törnqvist index. These are superlative index number formulae that can approximate a cost-of-living index. The difference in the 12 month growth between the CPI (locally collected data) and both the calculated Fisher and Törnqvist indices is on average 0.46 percentage points. However, this difference changes over time; during the two year period where 12 month growths could be calculated (2008 to 2009), the difference was seen to range from 0.32 to 0.62 percentage points for the Fisher and from 0.32 to 0.60 percentage points for the Törnqvist. The Törnqvist index and the Fisher index very closely follow one another.

One of the key observations of this analysis is that the geometric Laspeyres was found to give a close approximation to the Törnqvist index (and hence the Fisher). The difference between the Törnqvist and the geometric Laspeyres, in terms of 12 month growths, was found to be 0.002 on average over the period 2008 to 2009 inclusive. Like the Törnqvist and the Fisher however, this calculation of the geometric Laspeyres relies on the availability of timely weights. The geometric Laspeyres is potentially an attractive choice of index number formula as firstly, it is a weighted geometric mean, so could be considered consistent with the use of a Jevons at the elementary aggregate level (see for example Armknecht & Silver 2012); secondly it does not suffer from chain drift.

Two geometric variants of the CPI were calculated, both of which could be calculated in a timely manner. The first version was a geometric Lowe, as the weights used were those that are used in the CPI, meaning that this includes price updating in the calculation of the class weights. The second version was a geometric Young index, which used the same weights as the CPI but with price updating removed.

The geometric Lowe was found to be consistently above the Törnqvist index over the period analysed. The average difference in 12 month growth between the Törnqvist and the geometric Lowe was -0.04 percentage points, indicating that the geometric Lowe is in general close to the Fisher. Note however that Armknecht & Silver (2012) indicate that the geometric Lowe has little conceptual support. It could therefore be viewed as not being a practical alternative to a Lowe index.

The geometric Young was found to typically fall below the Törnqvist index over the period analysed. The average difference in 12 month growth between the Törnqvist and the geometric Young was 0.06 percentage points, indicating that in general the growth in the geometric Young is close to the growth in the Törnqvist index. This indicates that the use of this formula could lead to an index that is a closer approximation to the Törnqvist index than the current CPI. The geometric Young index can be calculated in “real-time” as there is no requirement for up-to-date weights. This is a result that has been seen in other empirical studies19.

The analysis was repeated one level below the “all-items CPI”, at the COICOP-12 level. The results at this level were mixed. Although the geometric Lowe appeared to closely approximate the Törnqvist index in some COICOP domains, this was not true for all domains. Analysing the difference between the geometric Lowe and the Törnqvist in terms of the ratio between the two index formulae following the approach used by Balk (2008) indicated that the difference can be

19 For example, Wood (2001), Armknecht & Silver (2012), van Kints & Bishop (2013)
related to the covariance between expenditure share changes and relative price changes. This was tested by calculating the correlation coefficient between these two terms. The differences in the correlations between COICOP-12 domains were seen to reflect how well the geometric Young approximates the Törnqvist. Therefore, although the geometric Young may not be a good approximation of the Törnqvist in all COICOP-12 domains, this can be understood in terms of the correlation between expenditure share changes and price changes.

Whilst the Fisher, Törnqvist and geometric Laspeyres formulae have been used in this analysis, the availability of only annual weights means that in practice the indices shown are not exact. They represent the only calculations that can be made given the data available at the time. In an effort to understand whether the conclusions of the analysis might change if monthly weights were available, a simulation study was undertaken. This calculated monthly weights by adding a normally distributed noise component to the annual weight at monthly intervals. The effect of using these pseudo-monthly weights was to add some more variability into the difference between the Törnqvist and the geometric Laspeyres; however these differences did not reach the levels of the difference between the Törnqvist and the CPI calculated using locally collected data.

The results in this analysis were also compared to similar analyses carried out in the literature. The difference between the CPI calculated using locally collected data and both the Fisher and the Törnqvist indices calculated here generally exceed the differences seen in similar work. However, there are differences in the applied methodologies that could lead to a difference in the results. It may be possible in future work to examine the effects of these different methodological assumptions, by applying these to the analysis and re-calculating the indices.

Further sensitivity analyses were carried out to ensure that the conclusions of this work were robust to the key methodological assumptions made. These included analysing the effect of assuming a constant basket of goods over the time period of the analysis and removing the rescaling applied to the weights.

Assuming a constant basket slightly lowers the Laspeyres and the geometric Laspeyres indices, but does not seem to affect the other formulae. Removing the rescaling of the weights affects all indices. However the effect of removing rescaling is to increase the level of all the indices by a similar amount, which will not affect the conclusions of the analysis.

This paper has presented the results of calculating superlative indices using data from the UK CPI. The 12-month growth in the CPI calculated using locally collected data exceeds a superlative index by 0.46 percentage points on average over the period 2007 to 2009. The geometric Laspeyres seems to provide a good approximation to the Törnqvist, but can still not be calculated in a timely manner. The geometric Young could however be calculated in a timely manner and the results indicate that this index approximates the Törnqvist relatively well. However, this conclusion does not hold for all COICOP-12 domains; further analysis would be required to understand whether the geometric Young index would be an appropriate choice.

Acknowledgments

The authors thank Natalie Weaden, Anselma Dobson-McKittrick, Mick Silver, Jim O'Donoghue, Joseph Winton and Rob O'Neill for their insightful comments on this work.
References


Boskin M.J., Dullberger E.R., Gordon R.J., Griliches Z., Jorgenson D.W, “Towards a more accurate measure of the cost of living”, Final report to the Senate Finance Committee from the Advisory Commission to study the Consumer Price Index


de Haan J., Balk B., Hansen C.B., 2009, “Retrospective approximations of superlative price indexes for years where expenditure data is unavailable”, paper presented to the Ottawa Group, available from: http://www.ottawagroup.org/Ottawa/ottawagroup.nsf/home/Papers+-+Authors+Last+Name


Van Kints, M., Bishop G., 2013, “Geometric Indexes and Substitution Bias in the CPI”, paper presented to the Ottawa Group, available from:


Annex A – Formula for chain linking

The double chain linking in the CPI can be expressed in the following way:

\[ P_{t,y} = P_{Dec, y-1}^{Dec, y-1} P_{Jan, Dec, y-1}^{Jan, Dec, y-1} P_{t, Jan, y}^{t, Jan, y} \]  

(A1)

where \( P_{t,y} \) is the chain linked index number in month \( t \) of year \( y \), \( P_{Dec, y-1}^{Dec, y-1} \) is the index number in December of the previous year \( y-1 \), \( P_{Jan, Dec, y-1}^{Jan, Dec, y-1} \) is the index number in January of year \( y \) based on December of year \( y-1 \) and \( P_{t, Jan, y}^{t, Jan, y} \) is the index number in month \( t \) of year \( y \) based on January of year \( y \). In practice, \( P_{Dec, y-1}^{Dec, y-1} \) refers to the last index number in the previously chained index series.
Annex B – Aggregate level formulae

The following formulae describe the aggregate level indices used in the analysis. Where relevant, these are given both in terms of both prices and quantities, but also in terms of lower level indices and weights.

The following notation is used throughout:

- \( p_i^t \) is the price of good \( i \) in period \( t \),
- \( p_i^0 \) is the price of good \( i \) in period \( 0 \),
- \( q_i^t \) is the quantity of good \( i \) in period \( t \),
- \( q_i^0 \) is the quantity of good \( i \) in period \( 0 \),
- \( N \) represents the number of goods,
- \( s_i^0 \) is the expenditure share for good \( i \) in period \( 0 \),
- \( s_i^t \) is the expenditure share for good \( i \) in period \( t \),
- \( s_i^b \) is the expenditure share for good \( i \) in period \( b \) and \( b \) typically precedes \( 0 \),
- \( s_i^{0b} \) is the price updated expenditure share for good \( i \) in period \( b \) using price changes between period \( b \) and period \( 0 \) (note that \( s_i^{0b} = s_i^b \frac{p_i^t}{p_i^0} \))

and

- \( I_{ij}^{0,t} \) represents lower level price indices (in practice these are elementary aggregate indices).

Laspeyres:

\[
P_{L}^{0,t} = \frac{\sum_{i=1}^{N} p_i^t q_i^0}{\sum_{i=1}^{N} p_i^0 q_i^0} = \sum_{i=1}^{N} s_i^0 I_{ij}^{0,t} \tag{C1}
\]

Paasche:

\[
P_{P}^{0,t} = \frac{\sum_{i=1}^{N} p_i^t q_i^t}{\sum_{i=1}^{N} p_i^0 q_i^t} = \frac{1}{\sum_{i=1}^{N} s_i^t I_{ij}^{0,t}} \tag{C2}
\]

Fisher:

\[
P_{F}^{0,t} = \sqrt{P_{L}^{0,t} \times P_{P}^{0,t}} \tag{C3}
\]
Calculating a Retrospective Superlative Consumer Prices Index for the UK

\[ P_T^{0,t} = \prod_{i=1}^{N} \left( \frac{p_i^t}{p_i^0} \right)^{(s_i^0 + s_i^t)/2} = \prod_{i=1}^{N} (I_i^{0,t})^{(s_i^0 + s_i^t)/2} \]  
\[ \text{(C4)} \]

Geometric Laspeyres:

\[ P_{GL}^{0,t} = \prod_{i=1}^{N} \left( \frac{p_i^t}{p_i^0} \right)^{x_i^t} = \prod_{i=1}^{N} (I_i^{0,t})^{x_i^t} \]  
\[ \text{(C5)} \]

Geometric Young:

\[ P_{GY}^{0,t} = \prod_{i=1}^{N} \left( \frac{p_i^t}{p_i^0} \right)^{x_i^h} = \prod_{i=1}^{N} (I_i^{0,t})^{x_i^h} \]  
\[ \text{(C6)} \]

Geometric Lowe:

\[ P_{GL}^{0,t} = \prod_{i=1}^{N} \left( \frac{p_i^t}{p_i^0} \right)^{s_i^h} = \prod_{i=1}^{N} (I_i^{0,t})^{s_i^h} \]  
\[ \text{(C7)} \]

Young:

\[ P_Y^{0,t} = \sum_{i=1}^{N} s_i^h \left( \frac{p_i^t}{p_i^0} \right) = \sum_{i=1}^{N} s_i^h I_i^{0,t} \]  
\[ \text{(C8)} \]

Lowe:

\[ P_{Lo}^{0,t} = \sum_{i=1}^{N} p_i^t q_i^h = \sum_{i=1}^{N} I_i^{0,t} s_i^{0h} \]  
\[ \text{(C9)} \]
Annex C – Combined weights

The following formulae were used to calculate a combined weight for the aggregation of the elementary aggregate indices to the all items CPI. The weight contains three components: the stratum weight, the item weight and the class weight. The weights are written in terms of the year to which the collected expenditure data refer and show the weights that would be applied to calculate a price index in year \( y \). The weights are written in terms of those applied in February to December and January with a base period of the previous December. This reflects the chain linking procedures that are applied. The weights summed to 1000 over all price quotes; the appropriate constant divisors are not shown below.

The notation is as follows: \( w_c \) denotes the class weights, \( w_i \) denotes the item weights and \( w_s \) denotes the stratum weights. The superscripts refer to the year the expenditure data were collected (\( y-2, y-1, \) and \( y \)) and the subscript \( pu \) indicates that a specific weight is price updated. As the item and stratum weights are collected across calendar years (July to June), the period of collection is indicated in the superscript by displaying the first year and the second year from which the expenditure information is collected. For example, \( w_{ij}^{y-2,y-1} \) would be the item weight calculated from expenditure collected from July in year \( y-2 \) to June in year \( y-1 \).

The weights for the Törnqvist index are calculated as the average of the weights used for the Laspeyres and Paasche indices as these most closely align with the base period and current period.

\[
\text{Formula} & \quad \text{January based on previous December} & \quad \text{February to December} \\
\text{Lowe (current CPI method)} & \quad w = w_{c,pu}^{y-2} w_i^{y-3,y-2} w_s^{y-3,y-2} & \quad w = w_{c,pu}^{y-2} w_i^{y-2,y-1} w_s^{y-2,y-1} \\
\text{Geometric Lowe} & \quad w = w_{c}^{y-3} w_i^{y-3,y-2} w_s^{y-3,y-2} & \quad w = w_{c}^{y-2} w_i^{y-2,y-1} w_s^{y-2,y-1} \\
\text{Young Geometric Young} & \quad w = w_{c}^{y-2} w_i^{y-2,y-1} w_s^{y-2,y-1} & \quad w = w_{c}^{y-1} w_i^{y-1,y} w_s^{y-1,y} \\
\text{Laspeyres Geometric Laspeyres} & \quad w = w_{c}^{y-1} w_i^{y-1,y} w_s^{y-1,y} & \quad w = w_{c}^{y} w_i^{y+1,y} w_s^{y+1,y} \\
\text{Paasche} & \quad w = w_{c}^{y-1} w_i^{y-1,y} w_s^{y-1,y} & \quad w = w_{c}^{y} w_i^{y+1,y} w_s^{y+1,y}
\]

For example, to calculate a price index in the months February to December 2009, the weights would be as follows:

\[
\text{Formula} & \quad \text{February to December} \\
\text{Lowe (current CPI method)} & \quad w = w_{c,pu}^{2007} w_i^{2007,2008} w_s^{2007,2008} \\
\text{Geometric Lowe} & \quad w = w_{c}^{2007} w_i^{2007,2008} w_s^{2007,2008} \\
\text{Young Geometric Young} & \quad w = w_{c}^{2008} w_i^{2008,2009} w_s^{2008,2009} \\
\text{Laspeyres Geometric Laspeyres} & \quad w = w_{c}^{2009} w_i^{2009,2010} w_s^{2009,2010} \\
\text{Paasche} & \quad w = w_{c}^{2009} w_i^{2009,2010} w_s^{2009,2010}
\]

Note that the calculation of the Paasche in 2009 requires class weights from 2009, which in practice are not available until 2011.
Annex D – Comparing the geometric Young to the Törnqvist

The approach taken by both Armknecht & Silver (2012) and Balk (2008) is to express the difference between geometric formulae as a logarithm of ratios. Taking this approach, the difference between the Törnqvist and the geometric Young can be written as follows:

\[ P_T^{0,t} = \prod_{i=1}^{N} \left( \frac{p_i^t}{p_i^0} \right)^{\frac{1}{2}(s_i^0 + s_i^t)} \]

Taking logarithms,

\[ \ln \frac{P_T^{0,t}}{P_GY^{0,t}} = \ln P_T^{0,t} - \ln P_GY^{0,t} \]

\[ \ln \frac{P_T^{0,t}}{P_GY^{0,t}} = \sum_{i=1}^{N} \left( \frac{1}{2} s_i^0 + s_i^t \right) \ln \left( \frac{p_i^t}{p_i^0} \right) - \sum_{i=1}^{N} s_i^0 \ln \left( \frac{p_i^t}{p_i^0} \right) \]

Using the approach of Balk (2008), this can be written as follows:

\[ \ln \frac{P_T^{0,t}}{P_GY^{0,t}} = \sum_{i=1}^{N} \left( \frac{1}{2} s_i^0 + \frac{1}{2} s_i^0 - s_i^b \right) \ln \left( \frac{p_i^t}{p_i^0} \right) \]

where \( p_i^{0,t} = \frac{p_i^t}{p_i^0} \) is the price relative for item \( i \).