Breaking Down the Differences between the CPI-U and C-CPI-U: Weights vs. Formula

Gregory Kurtzon
U.S. Bureau of Labor Statistics

All views expressed in this paper are those of the author and do not necessarily reflect the views or policies of the U.S. Bureau of Labor Statistics.
Overview

- A major question the price index literature is the appropriate formula to use for a price index
- Public concerns over what is being assumed
- For a major purpose of price indexes - the upper level aggregation of the U.S. CPI, – the formula has little relevance if current weights are used
- The Laspeyres-type CPI-U formula vs. the Tornqvist chained CPI-U (C-CPI-U) formula
- Using currently monthly weights in the CPI-U makes it a chained Laspeyres
- Theoretically, as chaining becomes more frequent, the formulas should approach each other and the current inflation rate
Spurious chain drift must be removed by using long term price relatives.

When this is done, the monthly chaining explains the large majority of the difference between the CPI-U and the C-CPI-U.

Using currently monthly weights holds quantities constant for only one month at a time instead of 36 months on average.

Other differences between the formulas make little difference over one month periods.

Therefore, monthly frequency is sufficient.
The CPI-U vs. C-CPI-U

- CPI-U is a Laspeyres-type index, cost of a fixed basket:

\[ CPI_{U,t-1,t} = \frac{\sum_i q_{iB}I_{it}}{\sum_i q_{iB}I_{i,t-1}} = \sum_i \frac{q_{iB}I_{i,t-1}}{\sum_i q_{iB}I_{i,t-1}} R_{it} \]

- \( q_{iB} \) denotes the implicit quantity for base period B for item-area i
- \( I_{it} \) is the index level, \( R_{it} \) is the index relative

- An arithmetic mean with shares updated by the index level

- C-CPI-U, Tornqvist formula, estimates a cost-of-living index (COLI)

- It’s a geometric mean with mean shares as weights:

\[ T_{t-1,t} = \prod_i R_{it}^{s_{it} + s_{i,t-1}} \]

- \( s_{it} \) is the expenditure share
In general, a COLI lower than cost of fixed basket due to consumer substitution
  - Obtain a higher standard of living when relative prices change

Tornqvist incorporates this with:

Assumed substitution: a geometric mean is lower than arithmetic mean

Current weights have direct information on substitution
  - Purchase less of higher inflation goods, reducing weight

C-CPI-U weights vs. CPI-U weights
  - From two months at a time instead of two years at a time
  - Updated every month instead of every two years
  - No lag instead of a year processing lag
Convergence of Indexes with Chaining

- Chained Laspeyres and chained Tornqvist are discrete approximations to continuous Divisia index:

\[
p_{t',t}^{div} = \exp \left( \int_{t'}^{t} \sum_{ia} s_{ia}(\tau) \frac{dlnp_{ia}(\tau)}{d\tau} d\tau \right)
\]

- The change in Divisia index = instantaneous, or ‘current’ COLI

- Therefore as chaining becomes more frequent, consumer substitution is incorporated by weights so that indexes converge to each other, the Divisia index, and yield the current COLI
Method

- Make intermediate indexes between CPI-U and C-CPI-U
- Change one part at a time that moves a cost of a fixed basket to a COLI
- Done in different orders to check robustness
- A. change weights first
- B. change formula first, to geometric mean
- C. change formula first, to Constant Quantity Tornqvist (CQTQ)
  - A first order approximation to a Laspeyres
  - Tornqvist formula that uses shares that hold the implicit quantities constant
Changing Weights First

- Spurious Chain drift: Can’t just use monthly updated weights in CPI-U
- Correlation between implicit quantity weights and relatives at monthly frequency
- Chained Laspeyres would fly up before falling toward Tornqvist
- Smooth relatives by using Long Term relatives:

\[ d_{it} = \prod_{\tau=1}^{t} \frac{1}{R_{i\tau}^{t}} = \left( \frac{I_{it}}{I_{i1}} \right)^{\frac{1}{t}} \]

- Works for this data because shares can be treated as constant for index construction
  - Geomeans was a good initial estimator of final C-CPI-U
  - No bias from smoothing high and low inflation periods
A. Change Weights First

1. Use LT relatives

From \( \sum_i \frac{q_{iB}l_{i,t-1}}{\sum_i q_{iB}l_{i,t-1}} R_{it} \) to \( \sum_{ia} \frac{q_{iB}d_{it}^\tau}{\sum_{ia} q_{iB}d_{it}^{\tau-1}} d_{it} \)

2. Use Tornqvist share weights

From \( \sum_{ia} \frac{q_{iB}d_{it}^\tau}{\sum_{ia} q_{iB}d_{it}^{\tau-1}} d_{it} \) to \( \sum_{ia} s_{it}^{TQ} d_{it} \)

3. Use geometric mean instead of arithmetic mean

From \( \sum_i s_{it}^{TQ} d_{it} \) to \( \prod_i d_{it}^{s_{it}^{TQ}} \)

4. Switch back to short term relatives

From \( \prod_i d_{it}^{s_{it}^{TQ}} \) to \( \prod_i R_{it}^{s_{it}^{TQ}} \)

5. Use non-interpolated relatives
B. Change Formula First, to Geomeans

- Chain drift also in geomean index
- 1. Use LT relatives
- 2. Use geometric mean instead of arithmetic mean

From \( \sum_{ia} \frac{q_{it}d_{it}^\tau}{\sum_{ia} q_{it}d_{it}^{\tau-1}} d_{it} \) to \( \prod_i \frac{q_{it}d_{it}^\tau}{\sum_{ia} q_{it}d_{it}^{\tau-1}} \)

- 3. Use Tornqvist share weights

From \( \prod_i d_{i\tau} \frac{\sum_{ia} q_{it}d_{it}^{\tau-1}}{\sum_{ia} q_{it}d_{it}^\tau} \) to \( \prod_i d_{i\tau}^{STQ} \)

- 4. Switch back to short term relatives
- 5. Use non-interpolated relatives
C. Change Formula First, to CQTQ

- No spurious drift problems or LT relatives needed, checks robustness of LT relatives

1. Use Constant Quantity Tornqvist

- From \( \sum_i \frac{q_i B I_{i,t-1}}{\sum_i q_i B I_{i,t-1}} R_{it} \) to \( \prod_i R_{it} \)

\[ L_{B,t} = \frac{\sum_i q_i B I_{it}}{\sum_i q_i B I_{i,B}} \]

2. Use Tornqvist weights

- From \( \prod_i R_{it} \) to \( \prod_i R_{it}^{S_{iTQ}} \)

3. Use interpolated weights
Decomposition Steps Changing Weights First

Index Level, Jan. '00 = 1

- CPI-U
- LT Lowe
- TQ weighted LT arithmetic index
- LT C-CPI-U
- C-CPI-U

Month, Jan.'00 - Dec. '09
## Results: Jan. ‘00 – Dec. ’09

Differences between index levels over entire period as fraction of the total CPI-U – C-CPI-U difference

<table>
<thead>
<tr>
<th>Replace ST Relatives with LT Relatives</th>
<th>Method</th>
<th>A. Weights First</th>
<th>B. Formula First: Geomeans</th>
<th>C. Formula First: CQTQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>8.38%</td>
<td>Same</td>
<td>NA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weight Effect: Replace Constant Quantity Shares with Actual Shares</th>
<th>Method</th>
<th>A. Weights First</th>
<th>B. Formula First: Geomeans</th>
<th>C. Formula First: CQTQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>85.98%</td>
<td>86.25%</td>
<td>96.22%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Formula Effect: Replace Arithmetic Mean with Geometric Mean</th>
<th>Method</th>
<th>A. Weights First</th>
<th>B. Formula First: Geomeans</th>
<th>C. Formula First: CQTQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2.05%</td>
<td>1.79%</td>
<td>1.45%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Replace LT Relatives with ST Relatives</th>
<th>Method</th>
<th>A. Weights First</th>
<th>B. Formula First: Geomeans</th>
<th>C. Formula First: CQTQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1.25%</td>
<td>Same</td>
<td>NA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Replace Non-interpolated Relatives with Interpolated</th>
<th>Method</th>
<th>A. Weights First</th>
<th>B. Formula First: Geomeans</th>
<th>C. Formula First: CQTQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2.34%</td>
<td>Same</td>
<td>Same</td>
</tr>
</tbody>
</table>
Different Base Periods

- Different start months used for time range studied
- Months after Dec. 05 not used since sample too small
## Results:

Mean Effects Across Initial Periods of each Step by Method of Breakdown as % of Total CPI-U vs. C-CPI-U Difference

<table>
<thead>
<tr>
<th>Method Description</th>
<th>Method A. Weights First</th>
<th>Method B. Formula First: Geomeans</th>
<th>Method C. Formula First: CQTQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replace ST Relatives with LT Relatives</td>
<td>6.29%</td>
<td>Same</td>
<td>NA</td>
</tr>
<tr>
<td>Weight Effect: Replace Constant Quantity Shares with Actual Shares</td>
<td>88.10%</td>
<td>88.33%</td>
<td>94.04%</td>
</tr>
<tr>
<td>Formula Effect: Replace Arithmetic Mean with Geometric Mean</td>
<td>3.64%</td>
<td>3.41%</td>
<td>2.82%</td>
</tr>
<tr>
<td>Replace LT Relatives with ST Relatives</td>
<td>-1.18%</td>
<td>Same</td>
<td>NA</td>
</tr>
<tr>
<td>Replace Non-interpolated Relatives with Interpolated</td>
<td>3.14%</td>
<td>Same</td>
<td>Same</td>
</tr>
</tbody>
</table>
Summary/Conclusions

- Using current weights and LT relatives is sufficient for CPI-U to approximate a COLI
- Formula is not relevant if chaining is frequent enough
- Monthly frequency is sufficient on the scale of the CPI-U vs. the C-CPI-U
- C-CPI-U does not depend on formula assumptions it is often criticized for
Contact Information

Gregory Kurtzon
kurtzon.gregory@bls.gov