Empirical Findings on Upper-level Aggregation Issues in the HICP
Thomas Knetsch / Deutsche Bundesbank
7 June 2022
1. Introduction

2. Methodology

3. Empirical results

4. Weight concept in benchmark index

5. Summary and conclusions

Underlying manuscript: Herzberg J., Knetsch T. A., Popova D., Schwind P. and Weinand S. (2022), Empirical Findings on Upper-level Aggregation Issues in the HICP.
1. Introduction
Motivation

• Inflation measurement is a relevant issue for monetary policy → e.g. ECB’s recent Monetary Policy Strategy Review

• Total HICP measurement bias plays a role for the derivation of the monetary policy target

• Various sources of inflation mismeasurement:
  (i) upper-level aggregation
  (ii) lower-level aggregation
  (iii) quality adjustment
  (iv) new products/new outlets
  (v) sampling
Contribution of this paper

• Focus on mismeasurement at the upper level of aggregation
• Distinguish between representativity and data vintage effects
  – representativity effect: choice of index formula
  – data vintage effect: reliability of weights
    → Assessing the trade-off that the use of more current weights may come at the cost of relying on preliminary data
    → Specific example: 2012 introduction of annual update of HICP weights with preliminary national accounts data on households’ consumption expenditures
(see companion paper Herzberg et al, RoIW, forthcoming)
Contribution of this paper

- Quantification of upper-level aggregation bias and uncertainty for the euro area HICP
  - analysis impossible for the euro area as a whole
  - Big-5 aggregate (Germany, France, Italy, Spain, Netherlands), representing more than 80% of euro area HICP

- By-product cross-country comparison
  → providing insights into (still) non-harmonised elements in HICP weight updating rules (to-price-update vs. not-to-price-update options)
Related literature

- Boskin Commission Report and the literature emerged worldwide in the aftermath of this famous study
- Greenlees/Williams (2010)
  → effect of shortening time interval for updating of weights
- Silver/Ioannidis (1994)
  → untimely weights, root mean squared error, European CPIs
- Herzberg et al. (forthcoming)
  → very similar evaluation framework
2. Methodology
Upper-level aggregation principles of HICP

• Laspeyres-type index

\[
P_{\text{HICP}}(y, m) = \sum_{i=1}^{I} w_{i}^{o}(y - 1, 12) \cdot \frac{p_{i}(y, m)}{p_{i}(y - 1, 12)},
\]

\[p_{i}(y, m)\] - price of good \(i\) in year \(y\) and month \(m\);
\[w_{i}^{o}(y - 1) \equiv w_{i}^{o}(y - 1, 12)\] - official HICP weight

• Annual updating of weights

\[w_{i}^{o}(y - 1) = \tilde{w}_{i}(y - \xi) \cdot \frac{c_{i}(y - 2; y - 1)}{c_{i}(y - \xi; y - 1)} \cdot \frac{p_{i}(y - 1)}{p_{i}(y - 2)} \cdot \frac{p_{i}(y - 1, 12)}{p_{i}(y - 1)}
\]

\[c_{i}(y; v)\] - households’ consumption expenditure of good \(i\) in year \(y\) as reported in the national accounts vintage released in year \(v\);
\[\tilde{w}_{i}(y - \xi)\] - (hypothetical) base weight referring to \(y - \xi\), \(\xi > 2\)

• To-price-update: \(\frac{p_{i}(y - 1)}{p_{i}(y - 2)}\) included

• Not-to-price-update: \(\frac{p_{i}(y - 1)}{p_{i}(y - 2)}\) removed, the Netherlands
Benchmark price index

- Superlative price index $\rightarrow$ Törnqvist formula
- Final national accounts (NA) weights [final vintage $v = \infty$]

\[
 w_i^f (y - 1) = \bar{w}_i (y - \xi) \cdot \frac{c_i (y - 1; \infty)}{c_i (y - \xi; \infty)} \cdot \frac{p_i (y - 1, 12)}{p_i (y - 1)}
\]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_0$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>2009</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2010</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2011</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2012</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2013</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2014</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2015</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2016</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2017</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2018</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2019</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>2020</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\circ$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
</tbody>
</table>
Bias and inaccuracy metrics

- Mismeasurement is quantified by $\frac{P^o_L}{P^f_{Tö}}$
  - $P^o_L$ Laspeyres-type index based on official weights (HICP)
  - $P^f_{Tö}$ Törnqvist index based on final NA weights (benchmark)
  - $P^o_{Tö}$ Törnqvist index based on official weights.

- Decomposition:

$$\frac{P^o_L}{P^f_{Tö}} = \frac{P^o_L}{P^o_{Tö}} \cdot \frac{P^o_{Tö}}{P^f_{Tö}}$$

representativity effect  data vintage effect
3. Empirical results
Time plots of monthly deviations

- Germany
- France
- Italy
- Spain
- Netherlands
- Big-5

Representativity effect
Data vintage effect
Total
Bias (=mean deviation)

\[
MD_{\text{Total}} = \frac{1}{T} \sum_{t=1}^{T} \ln \left( \frac{P_L^o(t)}{P_T^f(t)} \right).
\]

<table>
<thead>
<tr>
<th></th>
<th>Representativity</th>
<th>Data vintage</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>0.044</td>
<td>0.046</td>
<td>0.090</td>
</tr>
<tr>
<td>France</td>
<td>0.027</td>
<td>0.029</td>
<td>0.056</td>
</tr>
<tr>
<td>Italy</td>
<td>0.031</td>
<td>0.036</td>
<td>0.068</td>
</tr>
<tr>
<td>Spain</td>
<td>0.042</td>
<td>0.025</td>
<td>0.067</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.040</td>
<td>-0.069</td>
<td>-0.028</td>
</tr>
<tr>
<td>Big-5</td>
<td>0.037</td>
<td>0.030</td>
<td>0.066</td>
</tr>
<tr>
<td>Euro Area</td>
<td>0.022</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Uncertainty surrounding HICP inflation

- \( RMSD_{\text{Total}} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \ln \left( \frac{P_L^o(t)}{P_{Tö}^f(t)} \right)^2} \)
- IDR - Interdecile Range

<table>
<thead>
<tr>
<th></th>
<th>RMSD Representativity</th>
<th>RMSD Data vintage</th>
<th>Total</th>
<th>IDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>0.062</td>
<td>0.063</td>
<td>0.112</td>
<td>0.153</td>
</tr>
<tr>
<td>France</td>
<td>0.035</td>
<td>0.039</td>
<td>0.070</td>
<td>0.115</td>
</tr>
<tr>
<td>Italy</td>
<td>0.043</td>
<td>0.051</td>
<td>0.091</td>
<td>0.143</td>
</tr>
<tr>
<td>Spain</td>
<td>0.052</td>
<td>0.045</td>
<td>0.088</td>
<td>0.149</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.052</td>
<td>0.081</td>
<td>0.059</td>
<td>0.125</td>
</tr>
<tr>
<td>Big-5</td>
<td>0.043</td>
<td>0.036</td>
<td>0.075</td>
<td>0.097</td>
</tr>
<tr>
<td>Euro Area</td>
<td>0.028</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
4. Weight concepts in benchmark index
Final NA weights vs. full-information weights: Concepts

- Final NA weights keep construction principle of HICP weights and incorporate timely and more mature NA data.

- Full-information weights make complete use of the universe of information helpful for weight compilation, irrespective of when it becomes available.

  ⇒ crucial additional element: weights are compiled using information from all household budget survey (HBS) waves.

- Our view: Price index based on full-information weights is better proxy of “true” inflation than one based on final NA weights.
Final NA weights vs. full-information weights: Comparison

• final NA weights: + 0.05 pp
• full-information weights: + 0.07 pp

→ With final NA weights, only a lower bound of “true” data vintage effect can be approximated

Contribution of data vintage component to upper-level aggregation bias of German HICP
5. Summary and conclusions
Summary of results

- Total upper-level aggregation bias of the Big-5 aggregate (representing more than 80% of euro area HICP) falls short of one-tenth of a percentage point
  - Representativity and data vintage components contribute to overall bias in quite similar shares
- The interdecile range measuring the uncertainty surrounding HICP inflation due to upper-level aggregation is about one-tenth of a percentage point for the Big-5 aggregate
  - Wider interdecile ranges are observed for individual countries, suggesting that contrary developments tend to balance out in the aggregate
Conclusions

• Upper-level aggregation issues are one source of HICP mismeasurement
  – Results confirm the view that their contribution to overall mismeasurement is likely to be small

• To draw a full picture of upper-level aggregation issues, it is necessary to quantify the data vintage effect, in addition to the representativity effect
  – It is feasible to calculate final NA weights for the Big-5 aggregate. With this weight concept, however, it is possible to quantify the data vintage effect in terms of a lower bound

• Systematic cross-country differences in data vintage effects may be related to still non-harmonised elements of HICP weight updating rules
  – European price statisticians might think of future harmonisation