Empirical Findings on Upper-level Aggregation Issues in the HICP

Thomas Knetsch / Deutsche Bundesbank

7 June 2022

Table of contents

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

- Inflation measurement is a relevant issue for monetary policy
  $\rightarrow$ 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

- 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)

- 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
  $\rightarrow$ providing insights into (still) non-harmonised elements in HICP weight updating rules (to-price-update vs. not-to-price-update options)

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

2. Methodology

Upper-level aggregation principles of HICP

- Laspeyres-type index

$$P^o_{\text{HICP}}(y,m) = \sum_{i=1}^{I} w^o_i(y-1,12) \cdot \frac{p_i(y,m)}{p_i(y-1,12)} \ ,$$

$p_i(y,m)$ - price of good i in year y and month m;
$w^o_i(y-1) \equiv w^o_i(y-1,12)$ - official HICP weight

- Annual updating of weights

$$w^o_i(y-1) = \bar{w}_i(y-\xi) \cdot \frac{c_i(y-2;y-1)}{c_i(y-\xi;y-1)} \cdot \frac{p_i(y-1)}{p_i(y-2)} \cdot \frac{p_i(y-1,12)}{p_i(y-1)}$$

$c_i(y;v)$ - households' consumption expenditure of good i in year y as reported in the national accounts vintage released in year v;
$\bar{w}_i(y-\xi)$ - (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 → Törnqvist formula
- Final national accounts (NA) weights [final vintage $v = \infty$]

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

	vintage available by end of year										
reporting	block A		block B					block C			
period	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
y_0	×	×	×	×	×	×	×	×	×	×	×
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2009	×	×	×	×	×	×	×	×	×	×	×
2010	*o*	×	×	×	×	×	×	×	×	×	×
2011		*o*	×	×	×	×	×	×	×	×	*f*
2012			*o*	×	×	×	×	×	×	×	*f*
2013				*o*	×	×	×	×	×	×	*f*
2014					*o*	×	×	×	×	×	*f*
2015						*o*	×	×	×	×	*f*
2016							*o*	×	×	×	*f*
2017								*o*	×	×	*f*
2018									*o*	×	*f*
2019										×	*f*
2020											×

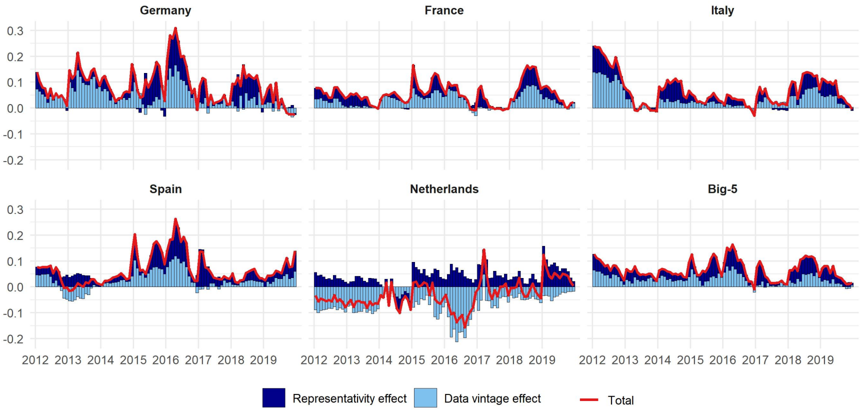
- Mismeasurement is quantified by $\frac{P_L^o}{P_{T\ddot{o}}^f}$
  - P_L^o Laspeyres-type index based on official weights (HICP)
  - $P_{T\ddot{o}}^f$ Törnqvist index based on final NA weights(benchmark)
  - $P_{T\ddot{o}}^o$ Törnqvist index based on official weights.
- Decomposition:

$$\frac{P_L^o}{P_{T\ddot{o}}^f} = \frac{P_L^o}{P_{T\ddot{o}}^o} \cdot \frac{P_{T\ddot{o}}^o}{P_{T\ddot{o}}^f}$$

representativity effect   data vintage effect

3. Empirical results

Time plots of monthly deviations



Bias (=mean deviation)

$$MD_{\text{Total}} = \frac{1}{T}\sum_{t=1}^{T} \ln\left(P_L^o(t)/P_{T\ddot{o}}^f(t)\right).$$

	Representativity	Data vintage	Total
Germany	0.044	0.046	0.090
France	0.027	0.029	0.056
Italy	0.031	0.036	0.068
Spain	0.042	0.025	0.067
Netherlands	0.040	-0.069	-0.028
Big-5	0.037	0.030	0.066
Euro Area	0.022	-	-

Uncertainty surrounding HICP inflation

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

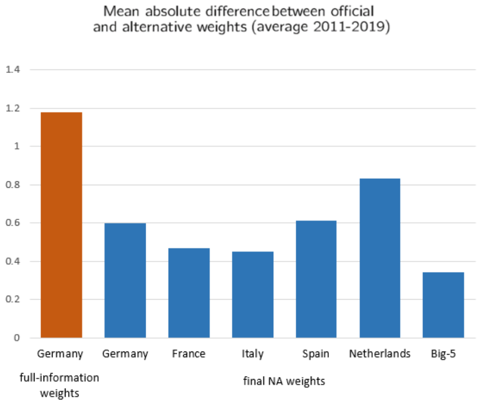
	RMSD			*IDR*
	Representativity	Data vintage	Total	
Germany	0.062	0.063	0.112	0.153
France	0.035	0.039	0.070	0.115
Italy	0.043	0.051	0.091	0.143
Spain	0.052	0.045	0.088	0.149
Netherlands	0.052	0.081	0.059	0.125
Big-5	0.043	0.036	0.075	0.097
Euro Area	0.028	-	-	-

4. Weight concepts in benchmark index

- 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

  $\Rightarrow$ 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



Contribution of data vintage component to upper-level aggregation bias of German HICP

- 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

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

- 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