



Multilateral Methods in the HICP

Vanda Guerreiro

Vanda-Maria.GUERREIRO@ec.europa.eu

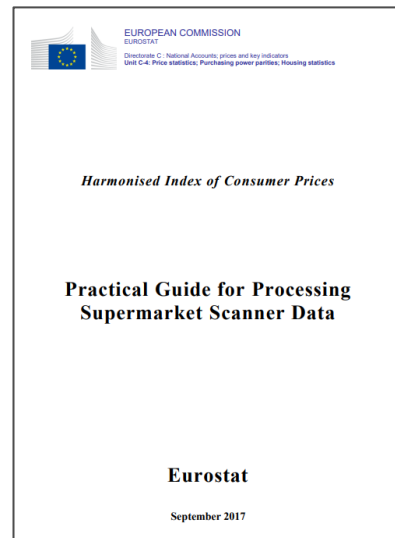
¹ Eurostat Unit C4

*Workshop on the utilisation and
integration of new data sources
for the Consumer Price Index
29-31 October 2024, EFTA House,
Brussels*

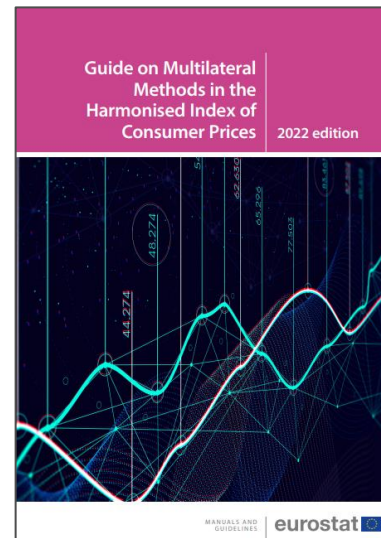
Overview

- What is scanner data ?
- Classification and Filters
- Index compilation:

Bilateral Methods



Multilateral Methods



What is scanner data ?

Variable description	Variable example
Start date of the week	02/feb/15
End date of the week	08/feb/15
Stock keeping unit	85235
GTIN code of the product	5449000000286
Brand description	Coca-Cola
Product type description	2L
Product info description	Regular (PET)
VAT-rate	6
Turnover	10000
Total quantities sold	4000
Sold by units or by weight	Units
Value belonging by unit or weight	2
Packaging description (Liter, Kilos, Pieces, ...)	L
Internal classification level 1	D
Internal classification level 2	E
Internal classification level 3	I
Internal classification level 4	K

Source: Statistics Belgium

Classification and Filters

Classification



Data should be classified according to ECOICOP or the lowest level of national classification



Initial step with each retailer:
classify all data

Item codes are partly unique for
every retailer



Afterwards: classify new products as part of the monthly production cycle

Classification

Matching

- Match an existing classified data set by GTIN and/or label

Use the internal classification of a retailer

- These data need to be available in the dataset
- Each retailer classifies its products in product groups
- Map this classification to the ECOICOP as accurately as possible
- Might not work for all groups: a retailer-specific group can be in 2 different ECOICOP groups

Classification

- Use machine learning
 - Use the data set of manually classified data
 - Different machine learning algorithms can be combined
 - Algorithm learns from classified data
 - Algorithm suggests code to be given to new data
 - The code should be verified
 - Algorithm learns again and becomes more refined

Filters

Outlier filter

Exclude observations with a high price change compared to the previous month (or any earlier period).

Dumping filter

Exclude observations if both price and quantity sold go down

Index compilation

Static and dynamic datasets

STATIC DATASET				
Product	Jan	Feb	Mar	Apr
Product 1	2.97	2.96	2.93	3.03
Product 2				
↓	3.64	3.5	3.36	3.37
Product 4				
Product 3	6.75	6.71	6.67	6.73

DYNAMIC DATASET				
Product	Jan	Feb	Mar	Apr
Product 1	2.97	2.96	2.93	3.03
Product 2	3.64	3.5	3.36	
Product 3	6.75	6.71	6.67	6.73
Product 4			3.37	3.37

Bilateral and multilateral indices

Bilateral price index	Multilateral price index
Measures price change	
From a base month	
To a current month	
Using only the base and current month	Using information from all months

Bilateral Methods

Bilateral method for dynamic samples

“The dynamic method automatically selects a representative sample of item codes for each consecutive set of two months (t and $t+1$, $t+1$ and $t+2$, $t+2$ and $t+3$ and so on) by selecting all matched item codes that have a turnover above a certain threshold and will include new and sufficiently important items whilst dropping items that are less important. The method resembles monthly replenishment and chaining.”

Eurostat: Practical Guide for Processing Supermarket Scanner Data

[Publications - Eurostat \(europa.eu\)](#)



PRACTICAL GUIDE | September 2017

Processing supermarket scanner data
English (1 MB - PDF)

Download

How to calculate the dynamic method

The sampled and imputed prices are used to calculate the index using a bilateral index such as Jevons

- Steps for calculating the dynamic method:
 - A dump filter
 - A low-sales filter
 - Imputations
 - Index calculation

Multilateral Methods

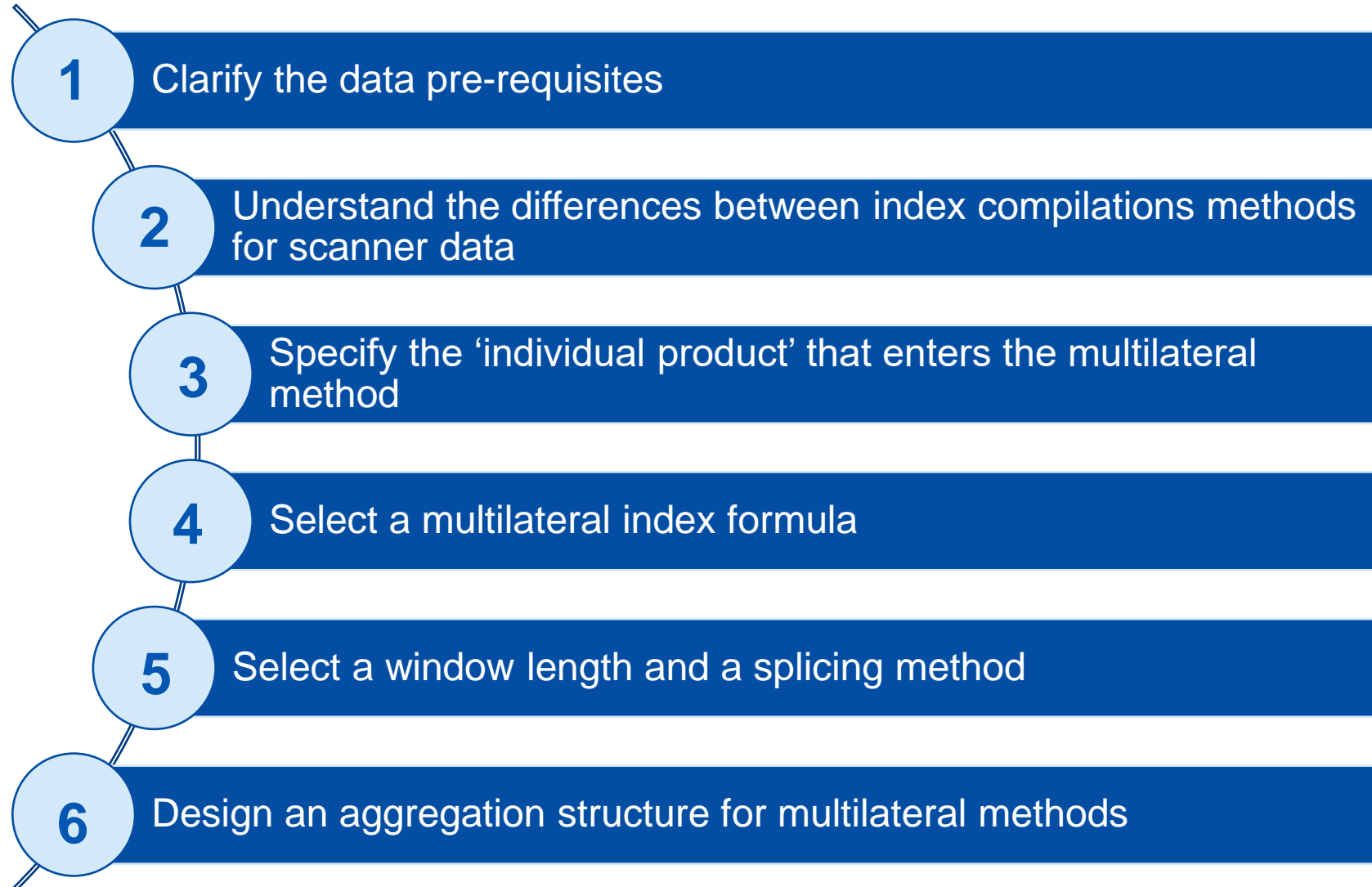
Multilateral methods

Guide on Multilateral
Methods in the
Harmonised Index of
Consumer Prices
2022 edition

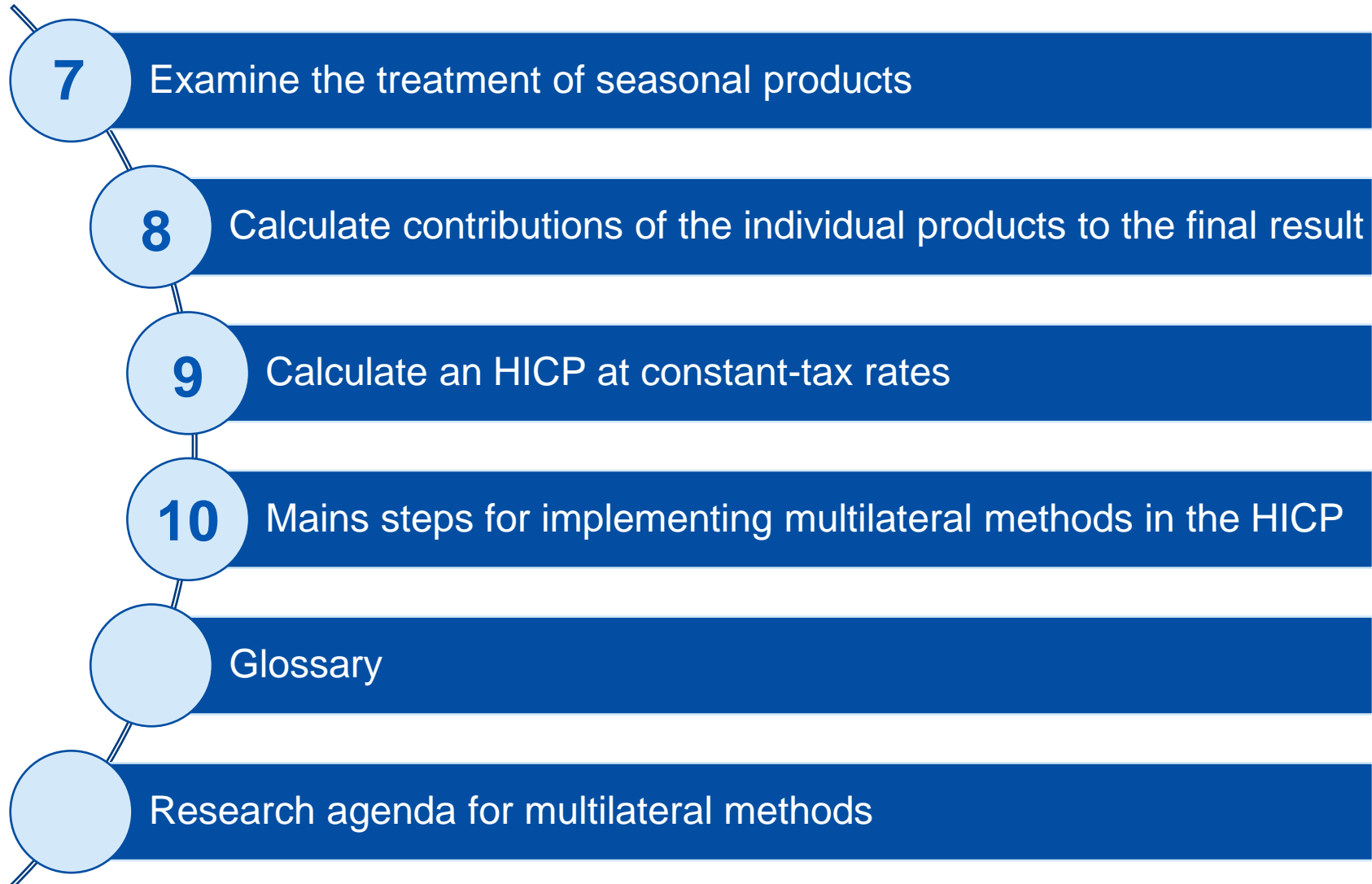
[Guide on multilateral methods in the Harmonised Index on Consumer Prices \(HICP\) — 2022 edition.](#)



Structure of the Guide



Structure of the Guide



Multilateral methods



Eurostat does not recommend a specific method.

The index formula for the compilation of the elementary aggregates must ensure transitivity and time reversibility.



The Guide addresses the most used multilateral methods.



The Guide will be updated on recent countries experience on methodological details and on the use of transaction data for other product groups than food and beverages.

Multilateral methods

GEKS

transforms many underlying bilateral indices between all periods into one transitive index.

$$I_{W(\text{GEKS}-Tq)}^{0,t} = \prod_{k \in W} \left(I_{Tq}^{0,k} \times I_{Tq}^{k,t} \right)^{\frac{1}{|W|}}$$

Weighted time – product dummy

a regression that includes dummy variables for the products and time periods that belong to the time window.

$$\ln p_i^t = \alpha + \sum_{\substack{r \in W \\ r \neq 0}} \delta^t D_r^t + \sum_{\substack{j \in N \\ j \neq 1}} \gamma_j K_j + \varepsilon_i^t$$

Geary-Khamis (GK)

can be seen as an extension of a Unit Value method with a quality correction.

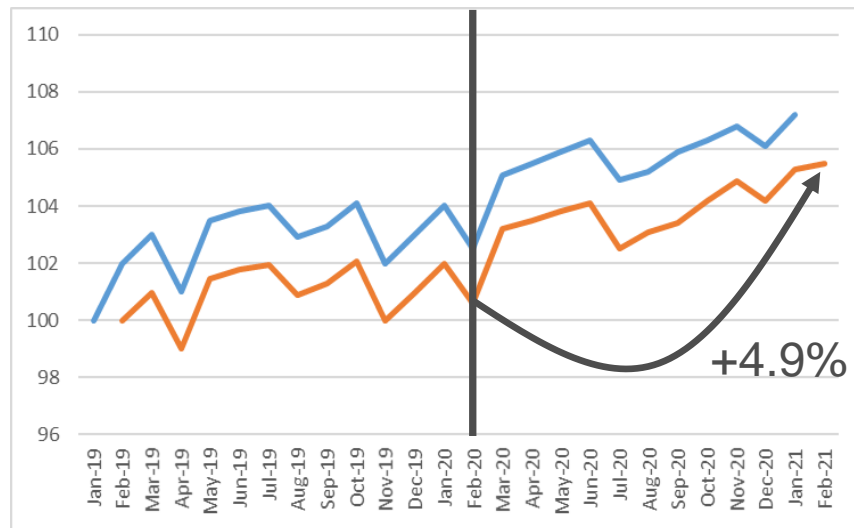
$$I_{W(\text{GK})}^{0,t} = \frac{\sum_{i \in N_t} p_i^t q_i^t / \sum_{i \in N_0} p_i^0 q_i^0}{\sum_{i \in N_t} v_i q_i^t / \sum_{i \in N_0} v_i q_i^0}$$

$$v_i = \sum_{z \in W} \frac{q_i^z}{\sum_{s \in W} q_i^s} \frac{p_i^z}{I_{W(\text{GK})}^{0,z}}$$

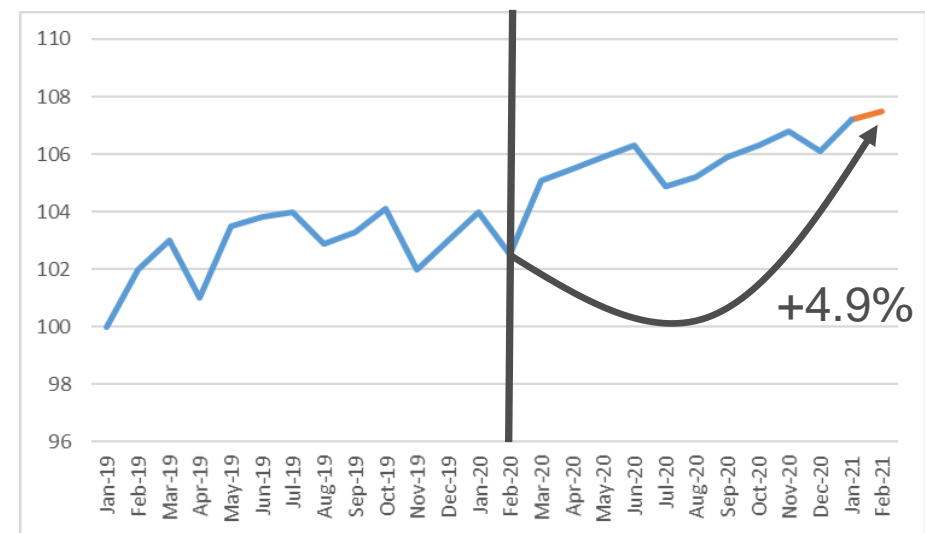
Window length and splicing methods

The 'half splice on published indices' method over a 25-month rolling time window (25-HASP).

The annual rate of change of the published index and of the most recently compiled multilateral index remains the same.



Index compiled over Jan19 to Jan21(blue) and over periods Feb19 to Feb 21 (orange).



Spliced (published) index from Jan19 to Feb21

Aggregation structures

At which level should the multilateral method be applied ?

Apply a multilateral method at a **more aggregated level:**

- cope with a dynamic product universe
- capture changes in the importance of individual products over time



Apply a multilateral method at a **more detailed level:**

- ensure consistency with higher-level aggregation
- obtain more stable results that are easier to analyse

Thank you