



The use of scanner data in the Luxembourg CPI/HICP: first lessons learned

Workshop on Scanner Data for HICP
26 – 27 September 2013 in Lisbon, Portugal

Claude Lamboray
claude.lamboray@statec.etat.lu



Project started in 2011

- Accessing scanner data
- Linking EAN codes to COICOP
- Compiling experimental price indices



Why scanner data?

Quality of the price indices

- Larger samples
- Better samples

Cost reduction

- Economies of scale
- Shift in resources (more IT)



Accessing scanner data

Bilateral meetings with the major national retail chains involving higher management

STATEC	Retail chain
Accurate inflation figures	Strategic value of scanner data
Personalized reports	Attitude of the other chains
Confidentiality	
Legal “obligation”	
Modernisation of public administration	
European project	Legitimate request by an NSI

Need to persevere

Maybe downgrade initial requirements to get started



Accessing scanner data

So far, 4 out of 6 retail chains have agreed to cooperate.

Written agreements

Automatic data transmission mechanism (e.g. FTP)

Main variables: turnover, quantities sold, label of the product, EAN code, shop classification



Linking EAN Codes to COICOP

In-house register that links each EAN code to a COICOP category (national 6 digit COICOP)

1. Mapping the shop classifications to COICOP
2. Dictionary with keywords
3. Manual classification

More than 90% of the EAN codes of COICOP Division 01 can be classified.



Linking EAN Codes to COLCOP

Maintenance of the correspondence table on a regular basis

Quality of the correspondence table (quality measures)

Temporal dimension

Chain specific codes



Compiling price indices

- Data cleaning
- Level of aggregation
- Price index formula
- Imputations

See also: CBS (2010). *The use of supermarket scanner data in the Dutch CPI.*
Paper written by Heymerik van der Grient and Jan de Haan.



Compiling price indices

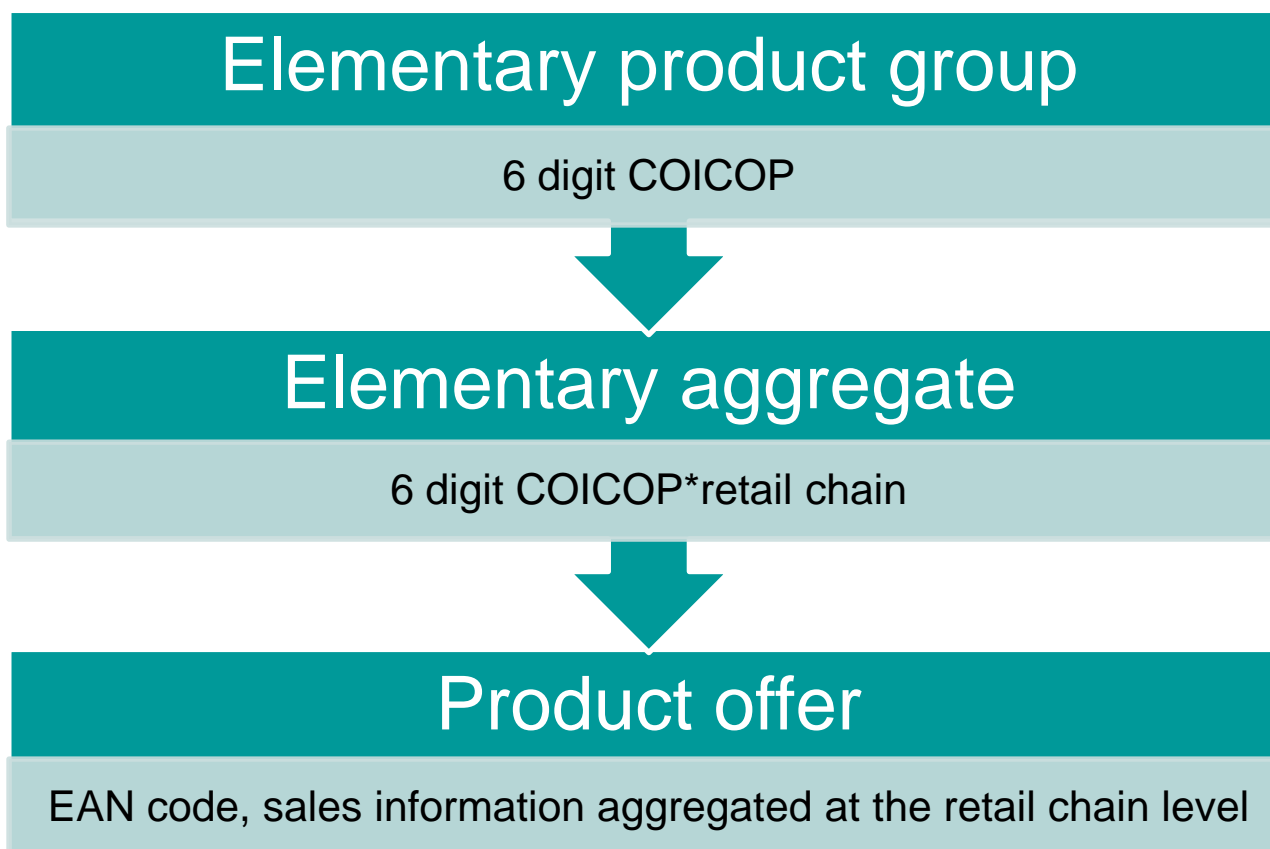
Data cleaning

- Excluding price increases of more than 300% and price decreases of less than -75%
- Excluding “contradictory” price and quantity movements: Prices and quantities increase (or decrease) by more than 50%



Compiling price indices

Level of aggregation





Compiling price indices

Price index formula

- Average share is computed for each product that is available both in the current and the preceding month.
- Products are sorted in descending order based on this average share.
- Products are selected such that the cumulated share exceeds 80% and that the selection covers at least 40% of the products.
- The price change between $t-1$ and t is an un-weighted geometric mean (Jevons price index) based only on the products that have been selected.
- These monthly indices are chained.



Compiling price indices

Imputations

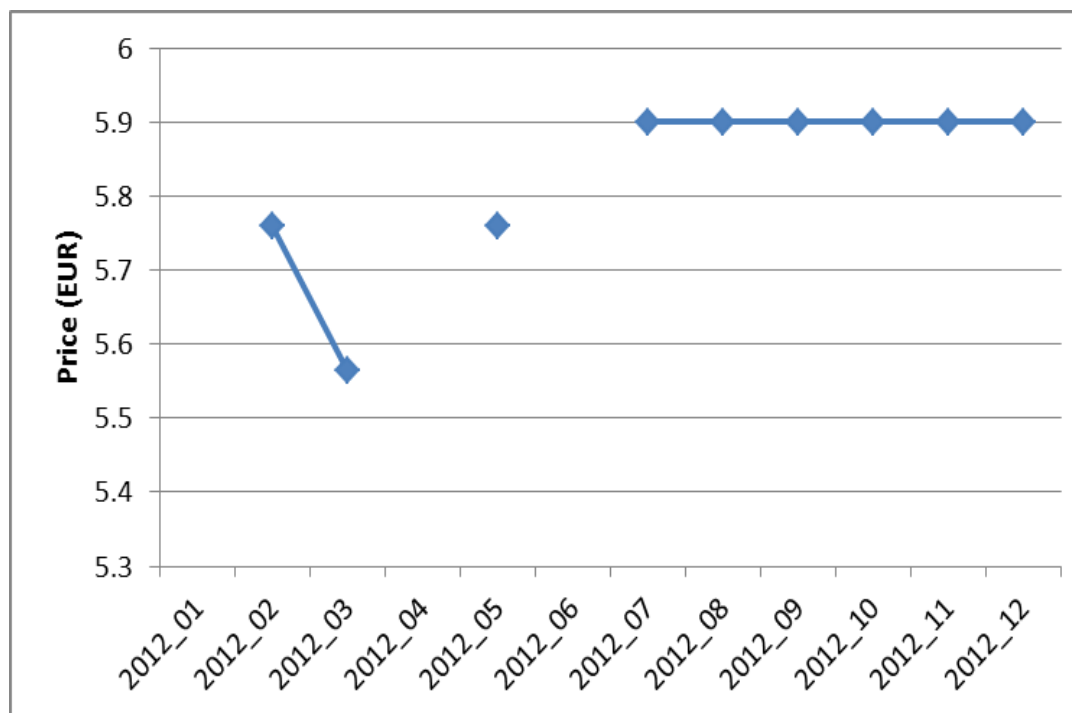
Prices can be temporarily missing:

- Out of stock
- No sales
- New EAN code
- Seasonal items



Compiling price indices

Imputations



Price for mineral water (24 bottles of 0,33L)



Conclusions

- Accessing scanner data is realistic.
- Linking EAN codes to COICOP seems feasible.
- The methodology for compiling price indices has to be further examined.
- Contacts with retailers that do not provide scanner data yet will continue.