Using transaction data in Consumer Price Index: Experiences at Statistics Netherlands

Antonio Chessa
CPI dept, Stat. NL (CBS)

US NAS Workshop: Scope, Components and Key Characteristics of a 21st Century Data Infrastructure
Session 2: Private sector data uses for national statistical purposes December 9, 2021

Outline

- Why transaction data?
- Accessing transaction data
- Current situation in CPI
- Implications for statistical process
- Final remarks

Scope of the CPI

- CPI used as primary measure of inflation
- Product categories: COICOP classification system
- Transactions: Business ↔ Consumer
- Types of CPI:
  - Constant tax: excludes tax and excise changes
  - CPI vs HICP: ‘citizens’ vs ‘domestic’

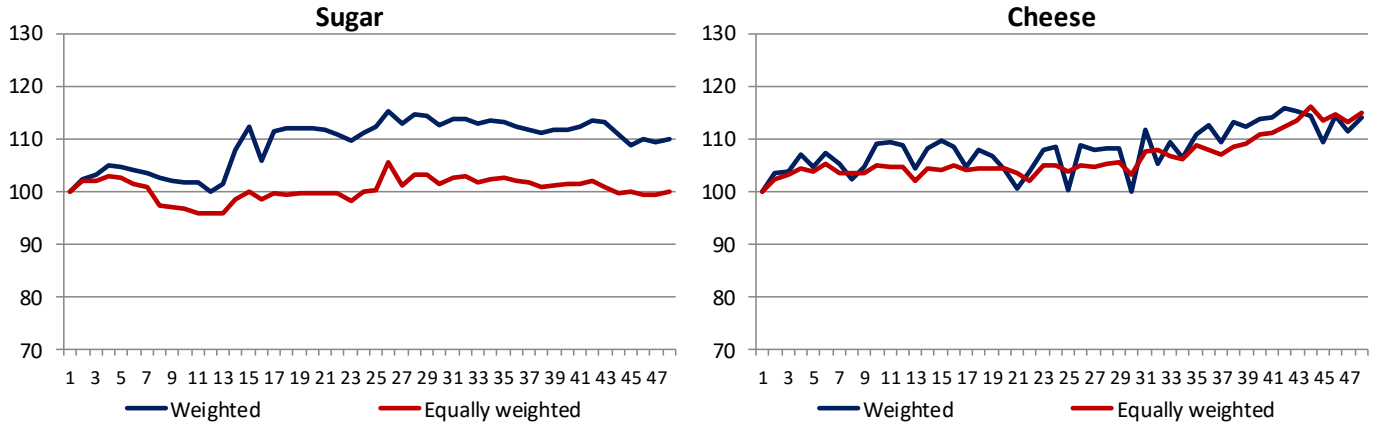
Transaction (scanner) data

- Data specified by GTIN (barcode)
- Data may contain **more than 100,000 items/GTINs** per retail chain
- **Expenditures** and **quantities sold** are available
- Prices are **transaction prices**, i.e. what consumers paid, on average, per item

Improving CPI quality with transaction data

- Much better product coverage
- From **static** to **dynamic** sets of products
- More sophisticated price index methods:
  - Product weights = expenditure shares
  - Weights can vary over products and over time
- Detailed statistics (week on week, by region, etc.)

Impact of product weights



Transaction data: The start in the 1990s

- Idea: Could we use transaction data for the CPI?
- A big retailer was contacted
- Research project also caught retailer's interest
- Project took 5 years
- Working towards implementation raised awareness about developing relationship CBS ↔ Retailer

Relations management (summary)

- Has a distinct role within our CPI team
- Team is responsible for the whole process, from first contact until contract arrangements
- Key aspects:
  - Explanation of importance of data
  - Confidentiality of data use
  - Security of data transmission
  - Content of data, timing and frequency of data delivery
  - Flexibility and return service of statistics office
  - National Law on Statistics 2003 (“CBS-wet”), includes scanner data

T data acquisition ‘roadmap’ (FYI) - 1

Aim: Extend the number of transaction data providers

1. **Market analyses**
   - Select the company: Relevant in turnover, relevant products, national coverage

2. **Analyze the company**
   - Existing contacts, current data arrangement
   - Arrange a visit and inform in advance

3. **Visit the company**
   - Contact with senior management
   - Discuss data needs for CPI
   - Explain role and position of statistics
   - Cost efficiency and less response burden
   - Inform on the legal obligation for data delivery
   - Discuss confidentiality and data security

T data acquisition ‘roadmap’ (FYI) - 2

4. **Start with test data and back data**
   - To check the completeness of the data
   - To test transmission and security line
   - To research methods for integration in CPI

5. **Formalize: Agree by a contract**
   - Frequency of data delivery
   - Variables and description, metadata, format
   - Technical arrangement of data transmission
   - Contact persons, etc..

6. **Start regular production for CPI**

7. **Provide companies price indices in comparison with total CPI**

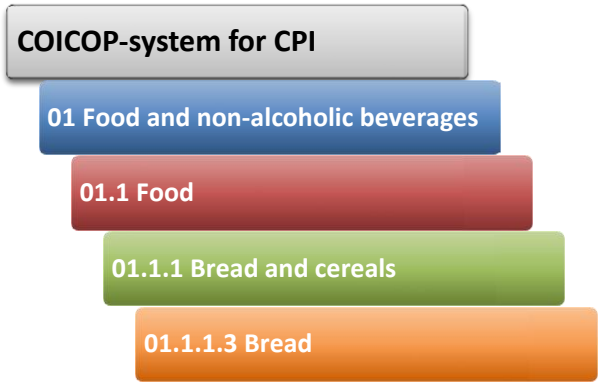
Flexibility towards retailer

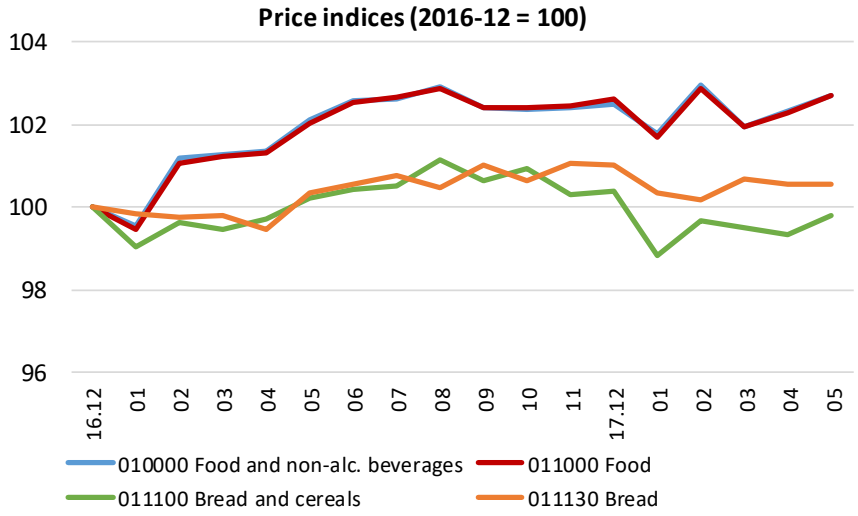
- Data delivery requires some (IT-)work initially
- We focus on limiting amount of work for retailers:
  - If retailers deliver data to other parties, we request a copy and try to limit modifications to the data
  - Retailers don’t have to deliver data for other statistics (e.g. retail statistics), as we take care of that
  - We handle classification of items/GTINs to COICOP

Our service to retailers

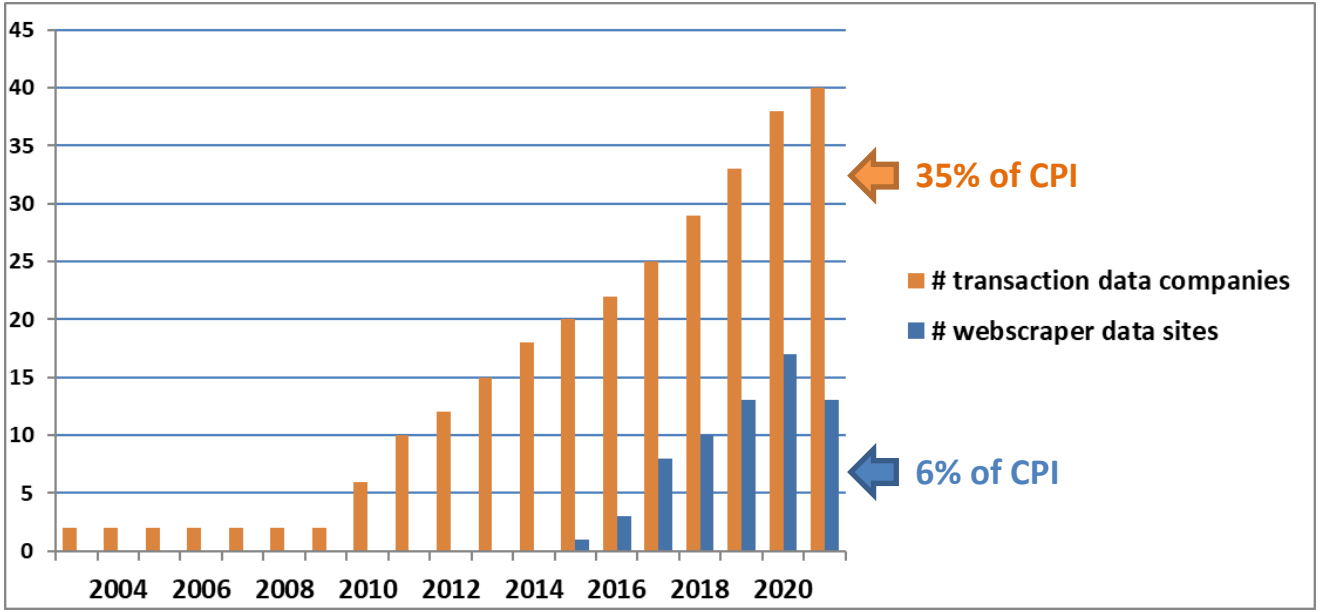
- CPI figures each month:
  - Retailer-specific figures vs whole market segment
  - Down to so-called COICOP-5 level
  - Retailer's index series are not shared with other retailers
- Retailers can always contact us for additional information and explanation

Example of CPI figures





Retailers by transaction and scraped data



Implications for statistical process

- Processing big data sets increases CPI complexity
- New methods for different stages:
  - Product definition/stratification [1]
  - CPI compilation is moving from traditional *bilateral* methods [2] towards *multilateral methods* [3]-[5]
  - Stage between CPI compilation and dissemination requires analytical methods and tools for index decomposition (product contributions)
- More automated processes

References included at the end

Final remarks

- Transaction data have big advantages:
  - Huge increase in product coverage
  - Higher quality inflation figures and more fine-grained statistics
- Data acquisition:
  - Retailers are generally willing to cooperate, understand importance
  - But willingness varies among retailers, negotiations can be lengthy
- Confidentiality and security are essential

References (1)

[1] Chessa, A.G. (2021). A Product Match Adjusted R Squared Method for Defining Products with Transaction Data. *Journal of Official Statistics*, **37** (2), 411-432. https://doi.org/10.2478/jos-2021-0018

[2] Eurostat (2017). Practical Guide for Processing Supermarket Scanner Data. https://circabc.europa.eu/sd/a/8e1333df-ca16-40fc-bc6a-1ce1be37247c/Practical-Guide-Supermarket-Scanner-Data-September-2017.pdf

[3] Chessa, A.G. (2016). A new methodology for processing scanner data in the CPI. *Eurostat Review on National Accounts and Macroeconomic Indicators*, issue 1, 49-69. https://ec.europa.eu/eurostat/cros/content/new-methodology-processing-scanner-data-dutch-cpi-antonio-g-chessa_en

[4] Chessa, A.G., Verburg, J., and Willenborg, L. (2017). A Comparison of Price Index Methods for Scanner Data. Paper presented at the 15th Meeting of the Ottawa Group on Price Indices, 1-12 May 2017, Eltville am Rhein, Germany. https://www.bundesbank.de/en/bundesbank/research/a-comparison-of-price-index-methods-for-scanner-data-635922

References (2)

[5] Chessa, A.G. (2021). Extension of Multilateral Index Series over Time: Analysis and Comparison of Methods. Paper presented at the *Meeting of the Group of Experts on Consumer Price Indices*, 2-10 June 2021 (online). https://unece.org/sites/default/files/2021-05/Session_1_Netherlands_Paper.pdf

Thank you!

ag.chessa@cbs.nl