Extending the Danish CPI with scanner data – A stepwise analysis

Jonas Mikkelsen
Statistics Denmark

Geneva, May/June 2012
Agenda

1. Motivation for the analysis
2. The main findings
3. Future analyses
4. Implementation
1. Motivation for the analysis

• To understand the **change of price concept** and the key differences between traditional price collection and scanner data.

• To investigate the **minimum necessary change needed** when implementing scanner data to our traditional production system. (Definition of the item basket)
  
  – Identification of methods that can **limit the amount of missing prices**.

  – Identification of methods that can **limit the amount of data in a manually maintained** item basket based on scanner data.
2. Main Findings – change of price concept

Turnover based prices from scanner data different from traditional collected shelf prices

Price = Weekly turnover per ean / Weekly volume ean = Weekly average unit price

The main difference is inclusion of volume discounts
Scanner data based CPI – Rice
Presumably a generic product and price stable category – Products chosen by DST

Current basket in scanner data vs. Published CPI

- The published CPI is based on single rice packages that is often not on discount
- The scanner data based CPI includes products bought on volume discount
2. Main Findings – definition of item basket

- Results that will be implemented:
  - Two weeks of data per month (is much better than one week)
  - Aggregating data to supermarket chain level
  - A turnover criteria securing that only important products are included
Scanner data based CPI – Minced beef
A product often bought on volume discount

SD turnover def basket (1w/m) vs. Published CPI

- Downwards bias (EANs leaving the basket on discount)
Scanner data based CPI – Minced beef
A product often bought on volume discount

SD turnover def basket, 2w/m data + chain aggregation vs. Published CPI

- If 2 weeks of data per month is used instead of 1 week per month the missing prices are limited and the downwards bias is limited
- If eans are also chain aggregated the missing prices are limited even more and the downwards trend is limited even further.
2. Main Findings – definition of item basket

- Results that leads to rethinking of basket definition criteria
  - The top three products per supermarket chain is often too restrictive
  - The 50% turnover share inclusion can be too restrictive
  - The 4 months in basket criteria needs further refinement
3. Future analyses

- Coicop specific criteria for item selection
  - Specific criteria were attrition of ean’s are high
  - Specific criteria for seasonal items

- The semi flexible basket system
  - Fixed basket as basis
  - Monthly maintenance compensating for:
    - Attrition (persistently missing in scanner data)
    - Products that have many missing prices through the year (stock outs)
    - Products that has become less representative over time
4. Implementation

• Test system with scanner data in production by 2013 analysing;
  – The production flow
  – Index behaviours
  – Comparison to current CPI

• Expecting to publish scanner data based CPI by 2014