Various data collection methods in the Norwegian CPI

Paper written by
Ingvild Johansen, Ragnhild Nygaard
Statistics Norway

Presentation at the joint UNECE/ILO Meeting on CPIs
Geneva 2012
Contents

• Our history of using paper questionnaires

• New data collection methods the last 10 years
  – Web questionnaires
  – Scanner data and data from secondary sources

• Efficiency gains from new data collection methods?

• Concluding remarks
The history of paper questionnaires

- The traditional data collection method in the Norwegian CPI

- Standardized questionnaires
  - Pre-defined representative items
  - Specific product description
  - Price
  - Check boxes for the identification of sales and quality differences and for the availability of the product

- The Norwegian Statistics Act of 16 June 1989
  - Collect information from businesses needed to provide official statistics
  - May impose compulsory fines
  - High and stable response rate for the Norwegian CPI
The history of paper questionnaires II

• Large response burden on businesses
  – About 33 minutes on average per month filling in food questionnaires (1995)
  – The average overall time was about 13 minutes

• Interviewers used for pre-collection visits
Motives for exploring greater variety of methods

• Focus on low response burden
  – High priority in Statistics Norway
  – But also pressure from businesses to deliver data electronically from the head offices

• Effectiveness

• Improving data quality

• Rapidly changing prices and more complex pricing structures

• Consumer behaviour is changing
  – The importance of the Internet

• Increasing availability of electronically data
New data collection methods

• The share of paper questionnaires in the CPI is declining
  – 10 years ago: nearly 40 per cent in terms of CPI weight share
  – Today: approx. 20 per cent
• While electronically reported data is increasing
  – Web questionnaires; 13 per cent
  – Scanner data; 20 per cent
  – Data from secondary sources; 10 per cent
Data collection methods based on CPI weight shares

- Paper questionnaires: 21%
- Web questionnaires: 10%
- Scanner data/electronically registered prices: 13%
- Computer assisted telephone interview: 13%
- Central data collection (Internet, telephone, e-mail, catalogues): 15%
- Other statistics in Statistics Norway: 6%
- Secondary sources: 10%
- No direct data collection: 2%
Web questionnaires

• System for electronic exchange with businesses (2004)
  – Possible to deliver the CPI questionnaires electronically
  – About 35 per cent choose web instead of paper questionnaire
  – Data also from local governments

• Improvements?
  – Many of the same challenges as paper questionnaires
  – Built-in validation checks
    ♦ Some of the editing process is transferred to the respondents
  – Increased flexibility
    ♦ Increased motivation among respondents to deliver accurate data?
  – Lower perceived burden
    ♦ Paper questionnaires may be perceived as old fashioned and ineffective
Scanner data

• The product codes are scanned into the cash registers of retail outlets when the items are purchased.

• The data is collected from the chains’ headquarters and contains information on price, quantity, type of outlet, location, period and description of the item.

• The items are identified by:
  – EAN (European Article Number), an international retail product code
  – PLU codes; internal chain specific codes
The use of scanner data in the Norwegian CPI

- First contact with retail chains: 1997
- Scanner data from one petrol station chain: 2003
- Full scale scanner data in CPI for food and non-alcoholic beverages: 2005
- Covering a broader range of products for medical products: 2012

Additional points:
- Full scale data collection from retail chains: 2001
- Scanner data from pharmacies: 2004
- Scanner data from the three largest pharmacy chains: 2011
Step-wise introduction of scanner data

• Initially we replaced prices from questionnaires with scanner data prices on predefined representative items
  – Still do that for non-food items

• Expanding the use of scanner data:
  – 2005: superlative price index for food and non-alcoholic beverages
  – 2012: broader range of products in price index for medical products
Data from secondary sources

• An alternative to scanner data is other types of electronic data from so-called secondary sources
  – Examples may be trade organisations and federal authorities

  – In the Norwegian CPI such examples are:
    ✷ The Norwegian Competition Authority
    ✷ The Norwegian Water Resources and Energy Directorate
    ✷ The Information Council for Road Traffic
    ✷ The Norwegian Booksellers Association
    ✷ The Consumer Council of Norway and the portal www.finansportalen.no
Scanner data and data from secondary sources – quality improvements?

• More control of the input data

• Actual transaction price for each product
  – The risk for manually reported errors connected to questionnaires is eliminated
    • Less dependent on the respondent’s motivation

• Larger data coverage
  – Full scale product coverage

• Opens up for possibilities for more advanced quality adjustment and calculation methods
Scanner data and data from secondary sources – quality improvements? II

• Increased response rate

• However,
  – Few data suppliers makes the statistical agency more vulnerable if they fail to deliver the data
  – Without the right tools, the amount of data may appear unmanageable
Efficiency gains?

• Large establishing costs for both supplier and recipient of the data
  – Testing
  – IT systems and technological solutions must be established
  – Written contracts

• Lower response burden on businesses
  – The average time per month filling in questionnaires has fallen from about 13 minutes to approximately 7 minutes

• Less use of interviewers

• The time spent in the statistical office producing CPI/HICP has increased by over 30 per cent the last 10 years
  – From about 8 full-time equivalents (FTEs) involved in the monthly CPI/HICP production and development in 2001 to 11 FTEs in 2011
Efficiency gains? II

• The increase in the number of FTEs probably a result of several factors
  – An increase in the number of official statistics within the Division
    ✷ Several indicators for underlying inflation, such as the CPI adjusted for tax changes and without energy products (CPI-ATE)
    ✷ The Rental Market Survey
  – Changes within the CPI production process caused by various data collection methods
    ✷ Less uniform production routines, vary between consumer groups
Concluding remarks

• Reduced response burden on businesses

• Improved CPI in terms of data accuracy and coverage

• Less uniform and more complex CPI production system