

15th Meeting of the Ottawa Group

10 – 12 May 2017

Session 8: Issues with new data sources

Unit values in scanner data – some operational issues

Jörgen Dalén, Consultant

The use of scanner data for CPIs involves a systematic use of unit values at the lowest level of aggregation. However, there are multiple ways of doing this with regard to the product, outlet and time dimensions. Both theoretical and practical concerns will influence research and development in this field based on data sets made available by the large retail chains.

This paper will discuss criteria for how to use scanner data, mainly for supermarket type data (food and daily necessities), where GTIN/EAN codes are used as product identifiers. Among special issues to be discussed in the paper are:

- An operational definition of homogeneity applied to product, outlet and time
- The relaunch problem (the fact that specified products tend to leave the market with a low price and an identical/very similar product entering the market at a higher price)
- The concept of a quality adjusted unit value index and its possible practical implementations

The paper builds on two reports on scanner data in 2014 and 2016 written within a project organised by ICON Institute in Germany and funded by Eurostat. Since this is a fast developing field, more recent developments may also be brought into the picture.