

Sampling design issues in Italian experience on scanner data and the possible integration with microdata coming from traditional data collection

Claudia De Vitiis¹

Scanner data for food and grocery in the modern distribution allow the use of sampling methods and index compilation using weights from quantities. This work focuses on a comparison among indices of elementary aggregates compiled using different sub-sets of series obtained through different selection scheme, in terms of elementary index bias and sampling error. This is a first experiment on a small data set related to some consumption segments and only permanent series. We compare probabilistic and non-probabilistic sampling selection schemes for different aggregation index formula. Sampling is carried out through cut-off selection of series based on thresholds of covered turnover and selection of 500 samples through pps sampling (using previous year turnover as measure of size). For each sample we compiled the elementary fixed base indices with three classical aggregation formulas: Jevons, Fisher and Lowe. Each estimate is compared with the corresponding universe value, evaluated on the complete set of permanent series.

¹ In collaboration with: C. Casciano, N. Cibella, A. Guandalini, F. Inglese, G. Seri, M. Terribili, F. Tiero