Emanuele Baldacci (Eurostat)
Ineke Stoop (The Netherlands Institute for Social Research)
Lilli Japec (Statistics Sweden)

Models of harmonization: now and in the future

Topic 7 – Better statistics for a globalised world

Keywords: comparability, quality, European statistics, benchmarks

Introduction

Cross-country comparability is particularly important for policy analysis use of statistics. As statistical data and indicators are increasingly requested to monitor policy implementation and assess its impacts, harmonized statistics are essential to analyse differences across countries. Comparable statistics ensure that no confounding factors linked to the way data have been collected, processed and disseminated get in the way of the cross-country comparison of policy effects or analysis of cross-country economic, social and environment trends. The issue of cross-country comparability is particularly important for European statistics, which are increasingly used to monitor progress in society and economy, as it is the case with the macroeconomic and social profile of countries, fiscal governance and Europe 2020 structural indicators. Important decisions are made at European level based on the statistical evidence provided by official statistics across different domains. While harmonization of official statistics to ensure cross-country comparability is a relevant quality dimension, there are however other equally important dimensions of quality that affect the usability of European statistics. These include relevance, accuracy and reliability, timeliness, clarity and accessibility. There may be trade offs between these dimensions of quality and comparability. Striking the right balance is therefore a complex issue.

Methods / Problem statement

Cross-country statistics' harmonization rely on a variety of methods and tools typically ranging from approaches that focus on inputs or target standardized outputs and more typically on a combination of these approaches (Koerner and Meyer, 2005). In particular, sources of variability of statistics across countries include the following: social and cultural environments affecting the way concepts are defined and measured data collection modes which affect respondents' attitudes and behaviours in surveys (De Leeuw, 2008) questionnaire design and structure having an impact on the way survey data and administrative data are collected and their conceptual value issues related to translation of terms in different languages and the associated possible lack of conceptual equivalence (Harkness, 2003) methodological (e.g., sampling frames, response rates, measurement errors, estimation methods) and financial resource differences in the collection and processing steps, which could limit what is feasible to achieve in terms of data quality access to data (Kolsrud et al., 2007). When data are collected through multiple sources, harmonization issues become more complex. Official statistics are increasingly relying on a combination of administrative and survey data adding potential unnecessary sources of variation in cross-country statistics.

Results / Proposed solution

The results of a review of European statistics practices shows that output harmonization methods are largely used in order to take into account country-specific conditions and minimize the financial burden of statistical production (Clemenceu and Museux, 2008). Input harmonization is more frequently used for

social surveys and in environmental statistics. Input harmonization approaches rely on standardized practices in each step of data collection from the sample design, to the questionnaire and data collection modalities, to data editing and imputation approaches (Ganda, Wolf, and Hadorn, 2010). The focus is on the conceptual equivalence of data collected and the harmonization of field work practices. In the European Statistical System (ESS) the ICT surveys and the Land Use and Cover Area survey (LUCAS) are examples of this approach based on a high degree of input harmonization. Outside the ESS, a good example of highly standardized work is the European Social Survey led by the Survey Research Centre, which aims at the highest possible harmonization in all steps of the survey life cycle: from questionnaire and translation to data collection work. Output harmonization practices are focused on concepts and definitions, methodologies for data processing and some degree of collection mode standardization. The EU Survey on Income and Living Conditions (EU-SILC) and the structural business statistics (SBS) are examples of this approach.

Conclusions

Cross-country statistics harmonization is important to ensure that data used for evidence related to policy are not biased by important factors associated to unnecessary variations in the way statistics are produced (Lynn, 2003). A variety of approaches is used in European statistics to ensure that this important quality dimension is fit for purpose, based on the use of the data and the users needs (Hoffmeyer-Zlotnik and Warner, 2013). Cross-country comparability is just one of the different dimensions of quality and as such trade offs could emerge between harmonization and precision or harmonization and timeliness, which need to be kept in mind. Notwithstanding efforts to improve cross-country harmonization practices, more work is needed to assess the degree to which differences across countries in various phases of the statistical business process have an impact on the usability of results. This includes reviewing harmonization practices, assessing impacts of methods and tools used in different stages of statistical production and benchmarking surveys (Stoop et al., 2010). This focus on comparability is more relevant now than in the past, as official statistics is moving away from reliance of sole survey data and administrative data are becoming increasingly used in statistical production. In the future, new sources of data, including sensor data, geopositioned data, and Internet-based data will increasingly be used to produce statistics (Citro, 2014).