ESSnet Project on Data Collection for Social Surveys using Mixed Modes (DCSS)

Annemieke Luiten[1], Karen Blanke[2]

[1]Statistics Netherlands, A.Luiten@CBS.nl

[2]Destatis, Germany, Karen.Blanke@destatis.ge

Abstract

The ESSnet Data Collection in Social Surveys has been initiated by Eurostat based on the fact that several NSIs have started projects to develop web-based data collection for social surveys. Eurostat's goal is to support early collaboration between countries. Apart from facilitating the introduction of web-based data collection for social surveys, the project takes mixed mode data collection into account. The exchange of experience and knowledge, sharing tools and identifying best practices are of major interest. The ESSnet is shared by a consortium of several countries: Partners are Statistical Offices of the Netherlands, United Kingdom, Finland and Germany. Three NSIs are acting as support-group-members: Denmark, Sweden and Italy. Destatis (Germany) acts as co-ordinator. Two external consultants are also involved in the project. The duration of the project is two years. A final public workshop on the results of the project is scheduled at the end of 2014.

Keywords: web questionnaires, CAWI, Labour Force Survey

1. Introduction

Appropriate and effective data collection methods are the basis for high quality data in the European Statistical System (ESS). However, there is much discussion at ESS concerning cost and time constraints which could potentially hinder quality aspects. Consequently, many countries are reconsidering their modes of data collection. Web-based data collection is cost-efficient and may counterbalance low response rates: specific groups of respondents may favour this form of data collection and modern technology provides various solutions for collecting data via Internet.

Therefore, Eurostat has initiated the ESSnet project "Data Collection in Social Surveys Using Mixed Modes" (DCSS) to support Member States in their development and implementation efforts. Eurostat is encouraging National Statistical Institutes (NSIs) to interact with and learn from each other in this consortium. Additionally, the findings of the project shall be disseminated to a broader audience.

The main objective of ESSnet is to stimulate early collaboration among countries, to harmonise and standardise national practices, to transfer knowledge, and to support countries that are less experienced with web data collection. To facilitate the introduction

of web-based data collection for social surveys, the project places specific emphasis on the impact of mixed mode aspects. These areas of interest are the two main work packages.

The ESSnet DCSS started in October 2012. The Consortium consists of five Statistical Offices: the Netherlands, Norway, United Kingdom, Finland and Germany. These partners will perform the main tasks and research. Three additional NSIs will act as support group members: Denmark, Sweden and Italy. These members shall participate at Consortium meetings when findings are discussed and further actions are defined. They will represent the perspectives of Member States and are highly interested in the issues. Two external experts, representing the European research community on mixed-mode methods, will assist as sounding boards. Within the Consortium, Destatis (Germany) acts as overall co-ordinator and project leader for the development of the web tool. CBS (Netherlands) leads the project on mixed mode data collection. The project will run for two years.

The project focuses on the Labour Force Survey (LFS). The LFS has been conducted in several modes throughout the ESS for a long time. The modes and their combinations differ among Member States. Most NSIs have adopted computer-assisted data collection modes for LFS (mainly CATI and CAPI). Until now, only the Netherlands applies a computer-assisted web questionnaire (CAWI) for regular data collection. However, several countries have started projects exploring the possibilities of using web-based data collection for social surveys as an additional mode. Consequently, there is a broad diversity of more or less experienced countries with regard to CAWI instruments. Against this background, this ESSnet project is an adequate starting point to harmonise and standardise a new instrument and share common features.

Important aims of this project are to identify the requirements of an appropriate web data collection tool (CAWI) for LFS and to examine the impact of introducing CAWI in a mixed mode LFS design. Although the focus is on LFS, the project is also of relevance for modernising the system of social statistics in ESS. The action addresses a number of important objectives set out in the "Wiesbaden Memorandum" endorsed by the Directors General Conference of the National Statistical Institutes (DGINS) held in September 2011: In the future, it will provide sound methodological data collection design of social statistics in order to maintain high quality standards while improving efficiency. Thus, it will contribute to a common European architecture for social statistics, and will lay the foundations for a higher degree of standardisation in the implementation of social surveys in the ESS.

2. The development of the web-based data collection tool

Whereas CAWI is quite common for business surveys, only a few countries have implemented CAWI for social surveys. Web questionnaires for social surveys require a different design in comparison to business surveys because respondents and topics differ. In addition, some business questionnaires are similar to forms or templates, rather than questionnaires. Respondents expect electronic questionnaires to use features they are

used to in market research or Internet programmes. Consequently, standardised CAWI applications used for business surveys cannot be applied directly as prototypes for social surveys. The CAWI instrument must be able to deal with error checks, have an automatic navigation feature, and collect data for all household members, among other requirements. Furthermore, it is paramount that the intent of every single question in each mode remains the same as in the basic mode applied (assurance of functional equivalence). Apart from the design and presentation of the questions itself, CAWI instruments have additional challenges to solve: Even though the functionality might be perfect, people do not necessarily know how to use the system. Therefore, the usability needs to be tested so that the quality of collected data is safeguarded. Finally, with regard to IT technology, the applications must be executable on the most important browsers and platforms.

In order to reach these objectives, several activities are planned:

(a) At the beginning, a review of existing CAWI instruments in social statistics will be made by launching a small survey of NSIs, with special focus on LFS. The review will consider issues like unimode versus specific mode questions, type of mixed mode combinations (e.g. CATI and CAWI, or PAP and CAWI), paging or screening designs, complexity of the questionnaires, offline or online technologies, software packages, or standardisation requirements which have to be adhered to by every NSI.

(b) Subsequently, all partners will conduct qualitative pretesting of their LFS CAWI applications. Partners will implement their pretests at different points in time; there will be a sequence of single pretests. The advantage of this approach is obvious: It is possible to share experience and define follow-up research questions, which might be covered and tested within the scope of another partner's pretest. Qualitative pretesting will be executed by cognitive interviews; three partners will also use eye-trackers. Features to be tested include login and password, navigation strategies, error checks, several household members in one CAWI instrument, presentation of instructions, error placement, error message design, prefilled questionnaires, don't know and empty options, functional equivalence of questions (unimode versus specific mode questions), mode-sensitive questions, and satisfying issues. Partners will have their specific emphases and country-specific challenges, but will also have areas of common interest.

3. Mixed mode data collection

Based on the Data Collection in Social Surveys Using Mixed Modes Workshop held September 22 - 23, 2011 in Luxembourg as well as Eurostat information and further discussion, the scope of the project has been extended to mixed mode data collection rather than focusing on web-data collection alone, since there will be several modes for data collection at the same time within countries and within the ESS. Web-based data collection (CAWI) is not (yet) suitable to be used as the only mode for official social statistics and will be used in combination with more traditional data collection modes, such as face-to-face or telephone modes. It is apparent that implementing mixed-mode

data collection in different combinations can be challenging and may impact technical issues, data collection procedures, fieldwork, weighting and estimation issues.

The second work package for mixed-mode data collection includes (a) a review of current surveying methods in ESS countries with specific emphasis on the LFS and (b) three major areas of research: organisation and case management systems for mixed-mode collection, measurement effects, and impacts on data processing and estimation.

The review shall comprehensively cover partially documented information regarding LFS in Member States: Based on a literature review and a small survey among all ESS countries, the combinations of data collection modes currently used in the ESS, as well as the sequences and order in which the different modes are implemented (e.g. concurrent vs. sequential), will be focused upon. The review will take into account quality requirements in the LFS and other pertinent issues which are covered by revision of the LFS's legal basis, expected to be adopted and introduced in 2016.

Based on the review on mixed mode data collection in the ESS, the partners will analyse the preconditions, advantages and pitfalls of mixed mode data collection use for LFS in their specific countries, with special emphasis on web-based data collection. Some partners in the project have access to large-scale databases, whereas others will introduce small-scale experiments. The research on mixed mode data collection will be conducted from three perspectives:

A. Organisation of mixed mode data collection
 The implementation of mixed mode data collection, and in particular of web-based data collection, will have important implications for various aspects of data collection organisation. It is of particular interest to understand which requirements the applied software package has to fulfil in a mixed mode environment. Using one single software package for several modes, at least at the national level, can reduce costs for IT development while at the same time reduce the problems associated with consolidating data collected via different modes into one uniform data model. The ESSnet partners will use at least two different software packages in their settings. Thereby, the consortium will have the opportunity to discuss drawbacks and determine the advantages of one or the other package.

In addition, under tight time constraints, as is the case for LFS, it is important to keep control of which respondent is contacted with which mode. If no response is received using one mode, no time should be lost assigning a different mode to this respondent. This is the responsibility of fieldwork management, for instance the coordination of field interviewers, the telephone studio, and the registration of incoming web responses (case management systems). Accordingly, it is not only essential that data sets to be collected are of the same format, but also that countries have suitable software to organise their data collection in an appropriate and efficient way.

B. Measurement issues

Data collection mode can have a substantial impact on statistical measurement, and hence the accuracy of the estimates. Although studies on mode effects have a long tradition in survey methodology, the definition, analysis and assessment of mode effects is far from being straightforward. First, mode effects differ with regard to the type of questions. For example, questions on opinions or attitudes are more sensitive with regard to mode effects than factual demographic variables, such as age or sex. Consequently, based on relevant literature and studies, mode effects will be studied in terms of core employment variables as well as different communication channels (e.g. visual, audible, oral / face-to-face, audio-visual) and the availability of an interviewer. Second, selection effects and how to distinguish between mode and selection effects are of major interest. Finally, the basic questions of which approach to implement for date collection, the unimode or the generalised approach,(mode -specific questions) in a mixed-mode environment for the LFS (unimode vs. functional equivalence) will be discussed. There are certainly pros and cons to both approaches. The key is to determine which approach best serves ESS and LFS conditions.

C. Data processing and estimation in mixed mode data collection

In addition to fieldwork organisation in general, as well as interviewer organisation and recruitment, response rates and panel attrition are affected. Different mixed mode designs will have different impacts on these aspects. Moreover, estimation and weighting procedures under mixed mode designs can evaluate mode effects, for example by adjusting for mode effects and selection biases in the data collection mode. Based on the research and experiences of partners, different approaches will be compared.

In contrast to the work package on web data collection tools, where all partners carry out qualitative pretests, partners in this work package focus on different aspects in order to cover the topics as broadly as possible, although main emphasis is on the assessment of mode and selection effects.

4.   Dissemination

Dissemination of knowledge is considered to be one of the main tasks to support the ESS. It encompasses sharing knowledge acquired in ESSnet, reporting results, and discussing recommendations with a broader audience beyond the consortium. This will be achieved by means of holding presentations at several workshops, conferences and working groups as well as establishing an ESSnet Portal on the Eurostat homepage. A final workshop on the main findings, which will be open for the research community, shall be scheduled at the end of the project.

In order to achieve broad sharing of knowledge and dissemination of findings, three main channels for dissemination shall be pursued:

A. ESSnet Portal
The project description, objectives, reviews, intermediate and final reports of the project will be presented on the ESSnet portal http://www.essnet-portal.eu. The project website will have various functions: It will serve as a dissemination platform and feedback forum, and will also provide contact to the experts working on ESSnet. Other non-participating countries will be invited to provide feedback and comments on the project, for example on projects with similar research questions. An invitation e-mail with the link to this website will be sent to the responsible institutions in all member states of the EU. The e-mail will inform about the project and welcome them to contribute their input regarding contents and proceedings.

B. Presentations at international conferences
The findings and results of this ESSnet will be presented at several workshops and/or conferences. The following conferences and workshops have already taken place and are scheduled for the respective presentations:

(a) **ISM (2012)**: At the Internet Survey Methodology Workshop (in Slovenia), basic design questions on CAWI instruments were discussed. Since the ISM is not limited to statisticians of NSIs, but also includes international researchers from universities and research institutes, the project was introduced with peers early on.

(b) **UN/ECE (2012):** The two work package leaders were invited to present their plans at the UN/ECE seminar 'New Frontiers for Statistical Data Collection' in Geneva, Switzerland, 31 October-2 November.

(c) **Eurostat**: Eurostat workshops, task forces and working groups related to LFS as well as the modernisation of social statistics will serve as a platform to present work in progress.

(d) **ESRA (2013)**: The next biennial conference of the European Survey Research Association will be carried out in 2013. By 2013, initial results of ESSnet can be presented and discussed.

(e) **Q2014**: At the end of the project period, the European Conference on Quality in Official Statistics will take place. An overview of the results and possible recommendations of the nearly completed ESSnet project will be presented. This conference also provides the opportunity to discuss strategic issues of mixed mode data collections, such as the impact of ESSnet on the social survey system in a broader sense.

C. Workshop on ESSnet DCSS
At the end of the ESSnet period, the findings of the project will be presented at a final workshop, organised and hosted by Destatis. ESSnet results and preliminary

recommendations will be presented. Non-participating Member States, other NSIs, and the international research community will be invited to attend the workshop and to discuss the topics covered by ESSnet.

5. Summary

The ESSnet project DCSS, is scheduled to cover two years and will last until September 2014. Coordinated by Destatis, a consortium of five partners (NL, NO, FI, UK, DE) and three support-group members (SE, DK, IT) will investigate a broad range of topics with regard to web questionnaires in particular and mixed mode data collection in general. The LFS will be used to analyse challenges and potential pitfalls. However, the project should also deliver results that will be applicable to the system of social statistics and to official statistics as a whole.

During the project, five qualitative pretests on CAWI instruments for LFS will be conducted by the different partners. Additionally, within the scope of mixed-mode data collection design, each partner will perform research on the organisation of mixed mode data collection, measurement issues (e.g. mode effects), data processing, and estimation by implementing mixed-mode surveys. In order to disseminate the project details and results, several presentations will be held at conferences, the ESSnet Portal will be established, and a final workshop will be scheduled. The project not only aims at offering guidance, but also investigates possible opportunities for sharing tools and knowledge, and identifying areas for common development. Moreover, it is of common interest for ESS to provide recommendations for web questionnaires and the survey design for implementing mixed-mode data collection. Since follow-up projects might be required in some areas, the project also serves to identify potential further actions.