**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

Seminar on Statistical Data Collection
(Geneva, Switzerland, 25-27 September 2013)

Topic (iv): Multiple modes of data collection

ESSNET PROJECT ON DATA COLLECTION FOR SOCIAL SURVEYS USING MULTIPLE MODES (DCSS)

Working Paper

Prepared by Karen Blanke, Federal Statistical Office Germany

I.   Introduction

1. The general purpose of the above project is to develop innovative tools for multiple mode and online data collection in social surveys in the European Statistical System (ESS). More specific it aims at compiling methodological knowledge and experiences regarding the impact on data quality while introducing a web mode within multiple mode data collection

2. The LFS was selected as an exemplary case, because this social survey is strongly harmonized and is already implemented using a multimode design with different modes being used in and across a large number of countries. The research results will therefore be highly relevant for all household surveys in the ESS. The project started in September 2012 and will take 24 months to complete.

3. The consortium led by Destatis consists of five partners ('Statistics Netherlands (CBS), United Kingdom Office for National Statistics (ONS), Statistics Finland (Stat. Fin.), Statistics Norway (SSB) and Federal Statistical Office Germany (Destatis), three support-group members (Statistics Sweden (SCB), Statistics Italy (ISTAT) and Statistics Denmark (DST) and two external consultants Peter Lynn (University of Essex) and Dirkjan Beukenhorst (former CBS).

II.   Two main work packages

4. **Work package: Web data collection** is led by Destatis. It is the ultimate goal to gather requirements of an appropriate web data collection tool (CAWI). Basically, there are two tasks to be covered:

- The starting point of the project was to review existing electronic web questionnaires in social surveys in the ESS, with particular emphasis on the LFS. Destatis carried out this review and will summarise the results for dissemination. Knowledge acquired during this review will be used to foster testing objectives, related to the both work package.

- All five partners conduct pretests on the newly or recently developed LFS web applications. Since all partners use different applications and software, testing is drawn on general design

features and specific issues (such as navigation, quantity, type and design of error checks), but also on IT-related topics (such as difficulties with different browsers or screen solutions, depending interviews, data security, confidentiality, system requirements, interfaces etc.).

5. The implementation of a CAWI instrument is a multi-dimensional challenge and needs sufficient evaluation, rather than simply programming the original questionnaire. The project will deliver recommendations on an appropriate electronic questionnaire design, so that other NSIs may implement electronic questionnaires based on recommendations and as a result be able to apply a harmonised, improved design.

6. **Work package: Multiple mode data collection.** Again, all five partners are involved. CBS Netherlands is leader of this WP. Within this work package the impacts of using mixed modes for the LFS in general and for implementing a CAWI instrument in particular will be analysed. As a first step, based on the query, NL will prepare a summary of the present state-of-the-art of mixed mode survey design in the LFS, based on the query and literature research. Furthermore, the partners will analyse the preconditions, advantages (mentioned before) and pitfalls of the use of multiple mode data collection in the LFS using the experiences and knowledge available in the countries of the consortium as well as relevant knowledge retrieved from either the query, research or carrying out experiments. The main areas in this work package are as follows:

   • the organization and case-management of multimode data collection

   • measurement issues (mode-effects)

   • data processing and estimation taking into account mode-effects and selection-effects

   • A key issue will be how to integrate web data collection in multiple mode data collection. Additional design factors need to be taken into account. It is recognised that NSIs apply different mode system concepts at present. As a result, they need to be aware that there is likely to be an effect on data quality as another mode is introduced and that the whole process of data collection needs to be revised. Thus topics are rather diverse. Each partner will tackle specific aspects and will distribute their work according to the expertise available in the consortium. One major task will be to collect the current knowledge on the impacts of mode effects in the LFS. The results of this investigation will finally be compiled and made available for the entire ESS. Moreover, first ideas on recommendations for good practices will be made and further methodological activities will be proposed to tackle the gaps in research and to contribute to a profound harmonisation of multiple data collection in social surveys of the European Statistical System.

III. Preliminary results

7. **Query** on 'Data Collection in Social Surveys'. Within the scope of both work packages, a query was implemented (field phase: March 2013 – May 2013). The main topics consisted in a) an overview of applied data collection designs in social surveys, b) a specification of the implemented mixed-mode data collection, c) main features of web questionnaire design, d) software package(s) used in the NSI-specific IT-environment and e) future prospects with regard to web data collection. All ESS countries were invited to participate as well as several overseas countries. Multiple reminders helped to raise the overall response rate to almost 85 % (data on core variables are available for 95% of the invited NSIs). Selected, preliminary findings: Currently, only two countries use a web questionnaire for the LFS in statistical production, whereas six countries are in the phase of testing a web questionnaire for LFS.

   • On the whole, web questionnaires are more common for the Census and the Household Budget Survey.

   • Within the next five years, more than one-third of the ESS countries are planning to use CAWI for social surveys.

- For the first wave of the LFS, CAPI is the major mode, whereas in the second wave, CATI becomes more relevant. Beyond that, there is an enormous diversity between the NSIs (e. g. Eastern and Western countries), using different modes within one wave and between waves.

- Mixed-mode data collection for the LFS is already reality for the majority of ESS countries. However, less than a third of the countries neither conducting mixed-mode data collection nor applying web questionnaires.

- The main reason for mixed-mode data collection: it is considered as cost-effective and respondent-friendly.

- With reference to the LFS, the implementation of a reference week is widely recognised as being an issue. The use of web questionnaires may require additional time for the data collection phase. Several countries clarified that solutions recommended as a result of the ESSnet must be in line with the provisions of the EU regulations (even distribution of interviews over the calendar year).

8. **Work package on web data collection:** All ESSnet DCSS partners carried out qualitative pretests according to the application and time schedule. The effort to conduct a pretest in each country is appropriate: Countries can confirm results on the same topics in or to complement each other by various topics or to cross-validate issues. A template on reporting pretesting results was provided by Destatis. It will be implemented as a common standard. Some preliminary, selected cross-country findings:

- *"Don't know" option, hard and soft checks:* "Don't know" has to be taken into account when allowing for proxy-interviews in official statistics, but quantitative research is required to find out how often it will actually be selected and whether it will rule out more substantial response options. Although being deeply anchored in the thinking of official statistics, ordinary respondents have problems in differentiating between hard and soft checks. Moreover, it needs to be discussed whether hard checks should be implemented only for mandatory surveys and for each question (vs. soft checks for voluntary surveys) and to which extent they should appear at the expense of respondents' motivation. There is no final recommendation so far.

- *Placement and design of explanations:* Data quality is at risk when respondents neglecting explanations since the LFS is partly based on rather complicated concepts, which are different from respondents' perspective. Best practice is to offer additional help directly below the *question*. Longer explanations need a specific button, hyperlink or roll-over function. The definitions should be short in wording, using bullet points rather than full-stop sentences.

- *Progress bars:* Fundamentally, progress bars are not recommended in complex surveys, since progress bars may become misleading, due to integrated skips.

- *Navigation:* Generally, respondents prefer a linear, automatic navigation through the web questionnaire. However, this design is challenging for NSIs to implement when applying the household approach (several members filling in one online questionnaire). Possible design solutions might be navigation trees or dashboard pages in order to give an overview on the structure of the questionnaire.

- *Individual vs. household approach:* The household approach may lead to a higher rate of proxy interviews (completing the questionnaire for other household members) . From a legal perspective, proxy-interviews would probably need the consent of the respondents which does not seem easily feasible in practice. Offering individual web questionnaires might be the better solution in this regard, but may extent the field period (waiting for individuals to enter data) and complicate the process of case management.

- *Dependent interviewing:* Even if respondents consider some prefilled information as useful, there might be concerns in terms of legal regulations and data quality (e. g. proxy interviews or faked interviews of former waves). Comparing different modes, dependent interviewing seems more

feasible when initial data collection is carried out by an interviewer-administered rather than a self-completed mode.

9. **Work package on multiple mode data collection:**

    - A lot of work has been dedicated to profound studies on concepts and definitions, before analysing data sets. A conceptual paper with regard to the definition, causes or nature of mode effects has been prepared and presented at the ESRA conference 2013. The international audience showed a positive reaction towards the project

    - Consequently, all partners conducted research on mode effects, either by data sets available due to former experiments within the scope of the LFS (the Netherlands and Germany) or by using data recently collected but by a different social survey (e. g. at Statistics Norway, United Kingdom, Finland). Finland will conduct a quantitative study in autumn by implementing a CAWI questionnaire for the LFS for the first time. The Netherlands, the only country within the consortium having a CAWI instrument in production will carry out several experiments.

    - Preliminary results by DE and NL indicated that mode effects in the LFS do not seem to be necessarily a major problem if they are taken into account carefully during questionnaire design and testing. This paved the way for a shift of attention to more organizational aspects of mixed-mode data collection that are equally important for less experienced NSIs in the ESS. Consequently, recent discussion, the project will address more emphasis on case-management systems and weighting. More attention is also placed on studying fieldwork strategies, besides research on mode effects.

IV.   Conclusion

10. The basic ideology behind the ESSnet DCSS project has ben derived from work by several NSIs within the ESS which have started to develop CAWI instruments with a view to integrating this additional mode in their already complex multimode survey designs. It is of particular importance to share experience and to identify an common harmonised way.In this situation Eurostat initiated the project. It is the purpose to make use of synergies by early collaboration and exchange of knowledge. This initiation of the project turned out to be almost perfect in time and contents: It helped either participating partners to go ahead or will repsond to several non-participating NSIs which are originally in phase of testing, experimenting or planning to introduce web questionnaires. Thus, first results showed already, that the ESSnet DCSS will provide support for NSIs that implement or plan to implement multimode and web data collection in social statistics so that they can build on previous methodological developments in other ESS member states and do not have to start from scratch in times of limited resources.

11. A final workshop, to be held in Wiesbaden, is scheduled for late summer 2014. Partners, support-group members, and consultants will be requested to contribute to an ESSnet-related topic. Furthermore, as part of the workshop programme, NSIs which are not part of the ESSnet DCSS will be encouraged to give presentations about their country-specific projects and research on the topic.

V.   References

The project website: http://www.cros-portal.eu/content/data-collection-social-surveys-using-multiple-modes-dcss.

Blanke, K.; Luiten, A (2012): "ESSnet Project on Data Collection for Social Surveys Using Multi Modes DCSS" The project was presented at the **UNECE Seminar** on "New Frontiers on Statistical Data Collection" in Geneva at the beginning of November 2012.
(www.unece.org/stats/documents/2012.10.coll.html.

Blanke, K. (2012): "ESSnet Project on Data Collection for Social Surveys Using Multi Modes DCSS" . Presentation at the annual **ESSnet Workshop** of Eurostat in Rome (http://www.cros-portal.eu/content/data-collection-social-surveys-using-multiple-modes-dcss).

Blanke, K.; Luiten, A. (2013): "ESSnet Project on Data Collection for Social Surveys Using Multi Modes DCSS": The project was posted at the **NTTS** (New Techniques and Technologies for Statistics) Conference in Brussels organized by Eurostat. http://www.cros-portal.eu/content/ntts-2013

Sattelberger, S. (2013):Setting further impulses for web household surveys in official statistics. Presentation at the conference of **CBS on 'Re-designing household surveys: sharing experiences'**. http://www.cros-portal.eu/content/data-collection-social-surveys-using-multiple-modes-dcss.

Sattelberger, S. (2013): ESSnet Data Collection in Social Surveys Using Multiple Modes. Brief presentation at the **Quest Workshop 2013** in Washington, DC. http://wwwn.cdc.gov/qbank/Quest/Quest.aspx

Luiten, A. (2013): Costs, timeliness, response and measurement errors: a review of multiple mode data collection in official statistics. Presentation at **European Survey Research Association** (ESRA) Conference 2013, Ljubljana. (http://www.europeansurveyresearch.org/conference).

Körner, T. (2013): Mode effects in the Labor Force Survey – Do they really matter? Presentation at **European Survey Research Association** (ESRA) Conference 2013, Ljubljana. (http://www.europeansurveyresearch.org/conference).

Gravem, D. (2013): Error prevention through interviewer evaluation. Presentation at **European Survey Research Association** (ESRA) Conference 2013, Ljubljana. (http://www.europeansurveyresearch.org/conference).

Pohjanpää, K; Järvensivu, M. (2013): The comparability o Don't Know answers between CATI and CAWI modes. Presentation at **European Survey Research Association** (ESRA) Conference 2013, Ljubljana. (http://www.europeansurveyresearch.org/conference).

Blanke, K. (2013): Web questionnaires in official statistics: Do's and don'ts. Presentation at **European Survey Research Association** (ESRA) Conference 2013, Ljubljana. (http://www.europeansurveyresearch.org/conference).