

TWINNING CONTRACT

JO/13/ENP/ST/23

Strengthening the capabilities of the Department of Statistics in Jordan



MISSION REPORT

on

Activity: 2.2: Design of surveys

Mission carried out by
Mr Mārtiņš Liberts, Central Statistical Bureau of Latvia
Mr Kai Lorentz, DeStatis

16–20 February 2014

Version: **Final**



Expert contact information

Dr Mārtiņš Liberts
Central Statistical Bureau of Latvia
Lāčplēša Street 1
Rīga, LV-1301
Latvia
Tel: +371 67 366 833
Email: martins.liberts@csb.gov.lv

Dr Kai Lorentz
Mathematical Statistical Methods
Destatis
Gustav-Stresemann-Ring 11
65189 Wiesbaden
Germany
Tel: +49-(0)611-75 25 89
Email: kai.lorentz@destatis.de

Table of contents

Executive Summary	4
1. General comments.....	5
2. Assessment and results.....	5
2.1. Household Surveys.....	6
2.2. Agriculture Surveys.....	7
2.3. Economic Surveys.....	8
3. Conclusions and recommendations	8
Annex 1. Terms of Reference	10
Annex 2. Persons met.....	12

List of Abbreviations

BC	Beneficiary country
CFS	Capital Formation Survey
CV	Coefficient of variation
DoS	Department of Statistics of Jordan
EU	European Union
GREG	Generalised regression
LFS	Labour Force Survey
PC	Personal computer
SPSS	IBM SPSS Statistics
RTA	Resident twinning adviser

Executive Summary

This was a successful project component activity. The aim of the activity was achieved through meetings and discussion with DoS employees, training sessions, and presentations. The main findings are that there is a room for improvements regarding the weighting of survey data and precision estimation has to be done more widely for all sample surveys. The main topics for the following activities will be imputation, weighting, and introduction with R software. Practical cases of the application of survey methodology will be discussed.

1. General comments

This mission report was prepared within the Twinning Project “Strengthening the capabilities of the Department of Statistics in Jordan”. It was the second mission to be devoted to the activity 2.2 “Design of surveys” within the component 2 “Sampling techniques” of the project.

The purposes of the mission were:

- To present and discuss how to design sample surveys;
- To present and discuss the choice of sampling methodologies;
- To present and discuss the state of the art in different surveys e.g. household, agriculture, business and economic statistics;
- Discussions of software solutions using SPSS;
- To present and discuss how to determine the sample size;
- To present and discuss weighting.

The consultants would like to express their thanks to all officials and individuals met for the kind support and valuable information which they received during the stay in Jordan and which highly facilitated the work of the consultants.

The views and observations stated in this report are those of the consultants and do not necessarily correspond to the views of EU, DoS, the Central Statistical Bureau of Latvia, or Destatis.

2. Assessment and results

The mission was carried out by the experts as a mixture of lessons (presentations) to DoS staff members and intensive discussions about sample surveys in the Kingdom of Jordan with specialists from subject matter statistic areas from DoS. The activities carried out were:

- Three meetings with Sampling Division (also covering household surveys), Agriculture Directorate, and Economic Directorate;
- Training session on sampling theory. See the attached files:
 - Sampling_part1.pdf;
 - Sampling_part2.pdf.
- Two training sessions on usage of IBM SPSS Statistics for sampling purposes. See the attached files:
 - MLiberts_2014_SPSS_Sampling.pdf;
 - SPSStraining.zip.
- Presentation session (about the Latvian and German experience in application of survey sampling methodology; particularly on sample size calculation and weighting). See the attached files:
 - MLiberts_2014_Sample_Size.pdf;
 - Sample_size_calc_ver021_ENG.xlsx;
 - Estimation_LFS.pdf.

It turned out, that knowledge of DoS staff members on sampling is to some extent well, but may be improved substantially. Especially the usage of state-of-the-art methods of weighting (like generalised regression estimation – GREG) should be explained in more details. Also imputation methodologies should be treated more extensively in one of the next missions of this activity. Last but not least, a repetition of basic concepts on sampling may be useful for a larger group of staff members from DoS.

Concerning concrete sampling plans shown to the experts, it seems that only in the area of household statistics standard errors of survey results are computed. This should be extended to all the subject matter areas.

Also IBM SPSS Statistics is used as the main statistical software for production; it would be good to have some other statistical software available for specific tasks in survey sampling (for example weighting, imputation, sample size calculation). The R software is a good candidate for such tasks. Basic training of R language could be planned for the next mission. Using R as an additional tool will not break the current production process of statistics. It can be used for some specific tasks independently. R code can be integrated within the SPSS procedures in the future to make the production process of statistics more automated. The DoS employees showed an interest in R software during the presentation of German and Latvian practices of weighting and sample size calculation.

2.1. Household Surveys

The questions of interest express during the meeting with Sampling Division were:

- Calculation of sample size;
- Sampling and weighting in general;
- The estimation of precision expressed as coefficients of variation (CV); using CV estimates for sample size calculation.

Two stage sampling is the usual sampling design used for the household surveys at the DoS. Stratified area sample is carried out at the first stage and dwelling sample is carried out for the second stage.

Possible under-coverage problem was detected for the first stage population frame. There are areas with new buildings which are not covered by the area frame used for the first stage sampling.

The IMPS software is used for precision estimation for household surveys. Several tables of the estimated CVs are added to the publications of household survey results.

The employees of the DoS are working on the implement of the survey on migration currently. The implementation of the survey has been consulted by the expert from the Eurostat.

The interest of having estimates about non-Jordanian employees (in context of Syrian crisis) was expressed. The figures about non-Jordanian employees are not published now from the Labour Force Survey (LFS); however the survey data about non-Jordanian employees are collected. The wish is to redesign the LFS to have better coverage of non-Jordanian employees. The population of the non-Jordanian employees is covered also by the Income and Expenditure Survey. But it is not obvious how to combine the results of both surveys.

The experts had a discussion about the Income and Expenditure Survey. CVs calculated by strata (sub-districts) were presented. The question about the sample allocation over strata was examined and recommendations regarding increase of sample size in some strata were given.

Other question of the discussion was the necessity of having at least two sampling units in each stratum. This is a minimal requirement that allows the derivation of variance estimates for the estimates of population parameters.

The sampling design for two stage sampling was another topic of the discussion as well as the possibility of having self-weighting sampling design in each stratum in case of two stage

sampling (by having equal size of sampling units at the second stage sampling). The sampling of large size primary sampling units was also discussed.

Important observation was that lack of the qualification of DoS employees is an obstacle to do more CV estimation (also for other sample surveys) and introduction of GREG estimation for survey data weighting.

2.2. Agriculture Surveys

The DoS currently works on introduction of sample survey for Jordan Valley. It was a census type survey for the last 10 years. There were financial support for that and need to survey all farms of the Jordan Valley completely.

The current sampling plan is to make a stratified systematic sampling. There are four strata (sub-districts). The order of farms is defined by the basin and the number of farm in each basin.

There was a discussion about the total sample size. Calculations using the frame of farms in Jordan Valley were done during the IBM SPSS Statistics training course. The expected CVs were calculated for sub-districts and total population using different sample sizes. The SPSS syntax (see in attached files) is provided to the employees of the Agriculture Directorate so they can use it for sample planning purposes. The calculated CVs for sub-districts and the total population with the total sample size 1000 farms are shown in Table 1¹.

Sub-district (strata)	Standard Deviation	Total Cultivated Area	Population size	Sample Size	CV
2	11.6	35 080	1 182	102	0.04
3	20.9	88 484	2 878	446	0.03
4	16.4	10 959	533	65	0.09
5	14.7	107 106	3 547	386	0.02
Total	-	241 627	8 140	1 000	0,02

Table 1: The sample allocation and expected CVs for the total cultivated area

There was a discussion about the Capital Formation Survey (CFS). Currently data for CFS is taken from six other surveys (Jordan valley (Aghwar), Uplands, Unorganized cattle, Sheep and goats, Organized cattle, Broilers, Layers, Mothers, Hatchers, beekeepers, Fish farms, All those surveys about Capital and investments in the agricultural sector: Buildings, animals, plantations, wells, water reservoirs, machineries and equipment, and all those survey are taken once a year in different times in the same year).

. Another topic of discussion was the possibility of combining survey data from different surveys. The aim of the DoS is to plan a new sample survey especially for the CFS. Development of the sampling plan for the CFS can be a topic for the next activities in the Twinning Project.

The employees of the Agriculture Directorate showed a great interest in the mission activities.

¹ The numbers of the table has to be recalculated using the most precise population frame.

2.3. Economic Surveys

There was one meeting with the representative from the Economic Directorate (Mr. Nimer Hashem Gharbia). The topics of the discussion were weighting, non-response adjustments, and imputation.

There was a discussion about the weighting in case of inactive enterprises in a sample and in the case when enterprise has changed the main activity. The current practice of weighting for the economic surveys could be not fully coherent with the general sampling methodology. The book of Sixten Lundström and Carl-Erik Särndal (2001) “Estimation in the Presence of Nonresponse and Frame Imperfections” was recommended as one of the references regarding the weighting in case of frame imperfections. Previously mentioned method has been used with regard to re-weighting and there was a significant gap in the results. Therefore, the current method was used after consulting sampling expert in DOS.

Also the issue of the imputation of large enterprises was treated. The first recommendation was to put as much efforts as possible to receive the data from the large enterprises. The current practice of imputation of large enterprises at the DoS (using the historical data adjusted for the general economic trends) is coherent with the general practice in EU.

Topic of discussion was also the Using statistical subdivisions (blocks) used in the population frame for economic surveys. There is not business register in Jordan. The attempt to build one has failed. This is an obstacle for DoS to keep the population frames for economic surveys up-to-date. DoS conducts an economic census every five years. This is a source for population frame. It is not updated between the census periods.

There was a proposal by Mr. Nimer Hashem Gharbia to follow-up the closed enterprises and enterprises under development after the census. That would allow partial update of the population frame between the census periods. The proposal was denied. We see that such approach is recommended. It would allow keeping the population frame more up-to-date which is important aspect regarding the precision of economic surveys. Especially it would help to treat the under-coverage problem (of new enterprises).

3. Conclusions and recommendations

The general recommendation for DoS:

- General capacity building regarding the theoretical and practical knowledge in survey methodology for the DoS.
- More extensive usage of survey and census data in survey planning process (especially in calculation of required sample size and optimal sample allocation);
- Usage of auxiliary information for weighting (for example GREG estimation);
- It is recommended to carry out the estimation of standard errors regularly for all the subject matter areas.

It is recommended to continue the work on capacity building of DoS staff members on survey sampling in one of the next missions on activity 2.5 or following activities. The recommended topics for the following activities are:

- Teaching of theory on sampling (especially on weighting, imputation, non-response treatment and sampling error estimation) to a broader group of statisticians from DoS;
- Introducing with the practical concepts of modern methods and tools on weighting (GREG estimation);
- Introducing with concepts and methods for imputation;
- Giving recommendations on specific practical topics proposed by DoS related to sampling methodology;
- Providing a training course with R software in linkage with previous mentioned topics.

The component leader expressed following recommendation for the activity 2.5:

- To cover unit and item imputation as one of the topics;
- To provide more practical knowledge. The expected proportion for theoretical and practical course material is 1:3;
- Introduction with R software.

The following actions are proposed to be done after the activity 2.2:

Action	Deadline	Responsible person
To provide the experts with practical problems faced in application of survey methodology. It should include detailed description of the problem. Possibly including appropriate data sets.	One month before the beginning of the activity 2.5.	The employees of DoS taking part in the activity 2.5.
<p>One of the PC classes at the DoS Training centre has to be prepared for the training session with R software:</p> <ul style="list-style-type: none"> • The internet connection for all PC is recommended. • The following software has to be installed on each PC: <ul style="list-style-type: none"> ○ R, available from: http://cran.rstudio.com/bin/windows/base/ ○ RStudio IDE, available from: http://www.rstudio.com/ide/download/desktop • A shared folder on the trainers' computer accessible from other PCs to distribute the files used during training. 	Before the beginning of the activity 2.5.	The responsible personal of the PC class at the Training centre.

Annex 1. Terms of Reference

EU Twinning Project JO/13/ENP/ST/23 16–20 February 2014

Component 2: Sampling techniques

Activity 2.2: Design of surveys

0. Mandatory results and benchmarks for the component

- Improve the capacity of DoS staff to understand and apply modern sampling techniques (Apr 2015)
- Assessment report on current situation (Jan 2014)
- Provide inputs to the design of surveys (Aug 2014)
- Conduct a training course in seasonal adjustment (Oct 2014)
- Give recommendations on how to deal with weights, imputation, non-response and sampling errors (Apr 2015)

1. Purpose of the activity

- To present and discuss how to design surveys
- To present and discuss the choice of sampling methodologies
- To present and discuss the state of the art in different surveys e.g. household, agriculture, business and economic statistics
- Discussions of software solutions using SPSS
- To present and discuss how to determine the sample size
- To present and discuss weighting

2. Expected output of the activity

- Knowledge gained on the state of the art of survey design
- Knowledge gained on software solutions using SPSS
- Transfer of the Latvian and German, and in general the European Union, experience in survey design
- Knowledge gained on how to determine sample size
- Knowledge gained on how to do weighting.

3. Participants

DoS

Mrs Ghaida Khasawneh, Head of Sampling Division (*Component Leader*)

Mr Mohammad Thnyyan, Head of Frames Division

Mr AbdAlnaserAljarere, Statistician in Sampling Division

Ms WafaaAmer, StatisticianAssistant in sampling Division

Mr AymanAlQasem, Statistician in Frames Division

MS experts

Mr Mārtiņš Liberts, Deputy Head of Mathematical Support Division, Central Statistical Bureau of Latvia

Mr Kai Lorentz, Head of Section, Mathematical-Statistical Methods, DeStatis

Other stakeholders taking part in the activity

Staff from other relevant DoS Directorates will be invited.

Programme for the mission

Time	Place	Event	Purpose / detail
Sunday, morning	Hotel / DoS	Meeting with RTA	To discuss the programme of the week
Sunday, morning	DoS	Meeting with BC Component Leader and BC Experts	Present and discuss how to design household surveys. Present and discuss the choice of sampling methodologies in household surveys.
Sunday, afternoon	DoS	Meeting with BC Component Leader and BC Experts	Present and discuss how to determine the sample size and weighting in household surveys.
Monday, morning	DoS	Meeting with BC Component Leader and BC Experts	Present and discuss how to design economic surveys, Present and discuss the choice of sampling methodologies, sample size and weighting in economic surveys.
Monday, afternoon	DoS	Meeting with BC Component Leader and BC Experts	Present and discuss how to design agriculture surveys, Present and discuss the choice of sampling methodologies, sample size and weighting in agriculture surveys.
Tuesday, morning	DoS	Meeting with BC Component Leader and BC Experts	Training of methods of sampling withdraws using SPSS for household surveys (Presentations and exercises).
Tuesday, afternoon	DoS	Meeting with BC Component Leader and BC Experts	Training of methods of sampling withdraws using SPSS for Economic and agricultural surveys (Presentations and exercises).
Wednesday, morning	DoS	Meeting with BC Component Leader and BC Experts	Presentations about the Latvian and German (in general the European Union) experience in survey design and sampling methodologies.
Wednesday, afternoon	DoS	Meeting with BC Component Leader and BC Experts	Presentations about the Latvian and German (in general the European Union) experience in determine the sample size and weighting in different surveys.
Thursday, morning	DoS	Meeting with BC Component Leader	Presentation of MS Experts' findings and agreement on the reached conclusions
Thursday, morning	DoS	Ad-hoc meetings	Final clarifications with BC Experts, preparation of report and presentation for BC Project Leader
Thursday, noon	DoS	Debriefing with BC Project Leader	Conclusions and decisions and their consequences for the next activity and the implied work programme for BC Experts

Annex 2. Persons met

DoS:

Mr AbedWadood Matouk, project leader

Methodologies and Statistical Techniques Directorate:

Mrs Ghaida Khasawneh, Head of Sampling Division (component leader)

Economic Directorate:

Mr Nimer Hashem Gharbia, Head of Industry and Energy Division

Agricultural Directorate:

Mr Bassam Zain, Head of Crops Production Division

Mr Mohammad Aljawarneh

Mr Abd Alnaser Obidat

Mr Haitham Ammari

Ms Eklas Shnikat

Ms Mai Sakarneh

Ms Suzan Al-Owaidy

Ms Hanan Alshabkeh

Ms Hend Alnaser

Ms Noura Abu Znemeh

Ms Sana Alaween

RTA Team:

Mr Thomas Olsen, resident twinning adviser (RTA)

Ms Amal Aliah, RTA assistant

Ms Deena Moghrabi, interpreter

Mr. Mohammad Junaidi, interpreter