

TWINNING CONTRACT

JO 21 ENI ST 01 22

Strengthening the capacity of Jordan's Department of Statistics in terms of compilation, analysis and reporting of statistical data in line with International and European best practices

MISSION REPORT

on

Component 2

Reviewing and developing the source data and methodology for producing Small Area Statistics

Activity 2.0

Assessment of current situation and development of the first rolling work plan for Small Area Statistics

> Mission carried out by Dr. Andrius Čiginas, Statistics Lithuania Dr. Michele d'Alo, The Italian National Institute of Statistics (ISTAT)

Amman, Jordan

October 17-20 October 2022

Version: 2.3









Authors' names, addresses, e-mails

Dr. Andrius Čiginas (MS Component Leader) Specialist at Data Technology Development Group Statistics Lithuania Gedimino Ave. 29 01500 Vilnius Lithuania Tel:+37065668785 Email: <u>Andrius.Ciginas@stat.gov.lt</u>

Dr. Michele D'Alo Head of a team in charge of small areas estimation and integration of data sources The Italian National Institute of Statistics (ISTAT) Via C. Balbo, 15 00184 Roma Italy Tel: +3946733026 Email: <u>dalo@istat.it</u>

Table of contents

Executive Summary	5
1. General comments	5
2. Assessment and results	5
3. Conclusions and recommendations	7
Annex 1. Terms of Reference	
Annex 2: Programme for the mission	15
Annex 3. Persons met	16

List of Abbreviations

- BC Beneficiary Country
- CoP General Population and Housing Census 2015
- DoS Department of Statistics
- HEIS Household Expenditure and Income Survey
- MS Member State
- PL Project Leader
- RTA Resident Twinning Advisor
- STE Short-term Expert
- SAE small area estimators
- SAS small area statistics

Executive Summary

The main goal of this first preliminary mission "Activity 2.0": within the component 2 of the twinning project JO 21 ENI ST 01 22 was the development of the first rolling plan of the future activities. The starting point for the development of a proper plan has been the assessment, inside DoS, of the knowhow on small area estimation methodology and the assessment of available sources of information to be used for computing small area statistics (SAS). In order to understand the state of the art and to share possible solutions based on the International and EU experiences on SAS, the first part of the mission was devoted to an overview of small area statistics in Jordan, including the state of the art, the possible needs and main challenges to cope with inside Dos and during the activities of the project, in terms of available methodologies, tools and information. In the following step, the presentations of the two EU experts involved in this first mission have had the aim to cover important issues related both to the relevance of the availability of further administrative information and steps needed to build an efficient process to be followed for computing small area statistics. The presentations have been followed by in-depth discussions on the methodological and IT needs and an agreement was reached on a list of necessary activities to be taken into account for the organization of the next missions and to set up the work plan.

1. General comments

This mission report was prepared within the Twinning Project "Strengthening the capacity of Jordan's Department of Statistics in terms of compilation, analysis and reporting of statistical data in line with International and European best practices". This was the first activity in component 2 on Small Area Statistics. The actions planned for this activity were carried out as scheduled.

The purposes of the mission were:

- Assessment of current situation for Small Area Statistics in Jordan
- Presentation of international and European standards
- Defining indicators
- Drafting the first rolling work plan for component 2 on Small Area Statistics

The consultants would like to express their sincere thanks to all officials and individuals met for the kind support and valuable information they received during their stay in Jordan and which greatly facilitated their work. The views and observations stated in this report are those of the consultants and do not necessarily correspond to the views of EU, Statistics Lithuania, or ISTAT.

2. Assessment and results

As stated in the Term of reference of the mission, the goal of this twinning project is the development of small area estimation methods (SAE) to be applied to some SDG indicators on poverty (GOAL1) and the possible extensions to other topics and indicators. Disaggregated statistics are required in order to allow more efficient policy on subgroups of the population that need specific development targets. Large-scale surveys are usually aimed at providing estimates of target indicators for the whole population, as well as for relevant subpopulations defined at the sampling stage. Design-consistent and design-unbiased direct estimates are produced for the parameters of interest. The sample size is not large enough to guarantee reliable estimates for all the target subpopulations that require more information than that usually collected through the survey. The small area estimation methods allow to overcome the lack of information in each small area of interest by combining surveys data with other types of information (e.g., administrative data or census data) that have wider coverage and allow gains in efficiency by borrowing strength also from other areas and/or other occasions of surveys. A small area estimator is applied whenever an indicator of interest can be computed from survey data and the sample size is not large enough to provide reliable estimates at the disaggregated level of interest if auxiliary data information correlated with the indicators of interest is available.

This first mission is aimed to define a rolling plan of activities to be carried out over the two-year life of the project, taking into account that the mandatory results of the component are:

- The assessment of administrative and other external data sources in the development of SAS and action plan developed.
- Development of methodology for producing SAS on poverty and provide recommendations on how this methodology can be applied to other areas within the DoS.
- Development training programs and manuals for use in the DoS and partner institutions based on pilot project outcomes.
- Implementation of communication strategy with stakeholders on strengthening small area statistics.

In this mission, we have focused on the first two points of the previous list. The first finding, regarding the availability of administrative data, is that it is necessary to await at least the first output of component 1 of the project that is directly involved in the identification and provision of the Jordan administrative data. Currently, there is no availability of any kind of administrative data, including demographic ones. Another aspect that bounds their actual use is the coverage of this information which seems to be limited only to the Jordan people.

With reference to the development of the methodology for producing SAS on poverty and on how this methodology can be applied to other areas within the DoS, the following precise steps would be applied. The first step, only partially accomplished in this mission, aims to clearly define the user needs in terms of both target indicators and the level of granularity desired. In this phase a clarification on all the available data that can be used in the estimation process should be also addressed, taking into account that the level of auxiliary information needed for a successful application of SAE methods depends on the type of target indicators (linear indicator are relatively easier to treat with respect to nonlinear ones). This phase is important for the second step which concerns the choice of the small area methodology that can be applied in each context, considering that some SAE methods are applicable to one form of indicators and different SAE methods might be necessary for other types of indicators. A third stage coincides with an important phase of quality evaluation of the whole process followed to compute a set of SAE estimates.

2.1 What have been achieved during the mission

Within DoS ways to produce small area statistics (SAS) have been experimented by the World Bank (WB). A poverty mapping methods proposed by Elbers C. et al (2003) has been applied, using the Household Expenditure and Income Survey (HEIS) data related to 2017/18 round and the 2015 Population and Housing Census data. The applied methodology and the results are explained in a methodological report (Céline Ferré and Matthew Wai Poi, 2020) provided to us by DoS. In short, the proposed methodology is based on a complex and specific unit-level model that cannot be straightforwardly extended to others types of indicators. Moreover, since covariates (auxiliary data, predictors) used for that modelling are derived only from the last available CoP data, the proposal of WB may be inefficient for off-Census years. The model is fitted using HEIS data that are representative at the governorate level over a full year. The sample is also representative at a national level, for every survey quarter. Therefore the national level and the 12 governorates are the geographic planned domains. In short, the ELL method proposed by the Word Bank fits a unit-level model on a log transformation of the HEIS's expenditure data using a set of auxiliary

information that is in common with the Census data set. The model parameters estimated with HEIS data are used to project values in the Population Census data. The estimates are computed at any level of interest using these predicted values. More details are provided by Céline Ferré and Matthew Wai Poi (2020).

The main pointed challenges during this mission are the lack of technical expertise within the staff of DoS on statistical modelling, SAE methods, and also the use of statistical software. SAE applications are moreover needed not only for poverty indicators. The lack of auxiliary information and its coverage for all the populations of interest is a gap that needs to be solved for more comprehensive use of SAE methods.

The success of SAE applications strongly depends on auxiliary information, which will be derived from administrative and other external data sources. Access to these sources depends on the results of component 1, and STEs were informed that some of the administrative data will be obtained within a realistic time frame for the project. On the other hand, some administrative data are already used in DoS but are not integrated into statistical surveys such as HEIS. For example, the Civil Status and Passports Department data are used in the production of population (demographic) statistics but are used neither to update the sampling frame nor to calibrate the sample weights to the demographic data. The latter innovations could lead to an improvement of the direct estimation at the governorate, district, and sub-district levels by applying calibration or generalized regression (GREG) estimators. Moreover, updated demographic data is a basis for SAE modeling.

A proper choice of statistical software is important for the project. For statistical calculations, DoS uses STATA and SPSS among others but, for this project and especially for component 2, R software would be more appropriate due to more possibilities for SAE and because of its general popularity in official statistics.

STEs delivered two presentations which may be annexed to this report. The component leader from Statistics Lithuania presented a very preliminary vision of SAE in Jordan based on the Twinning Fiche and other publicly available documents. The presentation of STE from ISTAT focused on international SAE implementation standards. All presentations were followed by discussions and the development of the plan for the next 5-6 of 12 missions.

Despite that extensive training is planned in sub-component 2.3, it was agreed that DoS staff should get hands-on experience on SAE as soon as possible. DoS plans to order introductory courses on R for that purpose.

3. Conclusions and recommendations

In relation to what emerged from the interesting information shared during the mission among the Resident Twinning Advisor (RTA) Charlotte Nielsen, the RTA Counterpart Eng. Mohammad Khalaf, the DoS experts and the two short-term experts, it was agreed to proceed with the definition of the activity plan that would give initially priorities on reducing the methodological expertise gap inside the DoS.

The first two missions are going to be devoted to the theoretical introduction to SAE methods and applications/illustrations of the theory using artificial data provided by MS.

In the third mission, HEIS survey data will be analysed together. The aim is to integrate administrative data used in DoS into the survey and thus improve the current (direct) estimation of poverty. This is also preparatory work before the first applications of SAE methods to real data.

The fourth and fifth missions are on the SAE applications to HEIS data, where the active participation of DoS staff is required as well. Here the data prepared in the third mission will be used. In addition, it

Strengthening the capacity of Jordan's Department of Statistics

is expected that some other administrative/external data will be already available and they can be included in the SAE modelling, too. To this aim, the fourth mission can consider a task to select relevant variables from the data obtained by component 1. For the estimation of poverty, administrative income data covering employee income, old-age benefits, and other benefits are desirable. To maintain the continuity of the recent WB project on the SAE application, new results will be compared with that of the WB under various scenarios.

Activity 2.1.1	Theory and best practice of Small area estimations – Part 1 Basic			
	smoothing and data manipulation			
Subject	 Definition and theory behind SAE methodology 			
	 Identification of the informative gaps with respect to the use small area 			
	methodology.			
	Clarification of the needs			
	• Basic smoothing - Production of direct estimate and their variances,			
	Production of other indirect estimators (synthetic and composite) from			
	designs-based perspective, Assessment – required threshold, etc.			
Methods	Morning – Theory and best practice – Presentations by MS			
	Afternoon - Practice and examples of artificial data provided by MS			
Resources	MS, BC			
Duration	4 Days			
Output	Common understanding of definitions and concepts			
	Best practice of basic smoothing introduced			
	Practice and experiences from artificial data obtained			
Time schedule	February 2023			

Activity 2.1.2	Theory and best practice of Small area estimations – Part 2 Model-			
	based Small area estimations			
Subject	Model-based small area estimations			
	 Definition and theory behind SAE models 			
	Characteristics of indicators			
	 Disaggregation to small domains 			
	• Assessment			
Methods	Morning – Theory and best practice – Presentations by MS			
	Afternoon - Practice and examples of artificial data provided by MS			
Resources	MS, BC			
Duration	4 Days			
Output	Common understanding of definitions			
	Best practice of model-based SAE introduced			
	Practice and experiences from artificial data obtained			
Time schedule	February 2023			

Activity 2.1.3	Assessing DoS data and updating census population, calibrating and creating covariates for SAE modeling		
Subject	 Experiences with SAE methods and work on real DoS data Assessing content, coverage and quality, and availability of DoS data for estimation of poverty Update census population Calibrate weights using the updated demographic information Prepare covariates for SAE modeling 		
Methods	Peer-to-peer work		
Resources	MS, BC		
Duration	4 Days		
Output	Improvement of the efficiency of direct estimates by means of better use of survey data and auxiliary information on population		
Time schedule	March 2023		

Activity 2.1.4	Applying theory and lessons learned on DoS data - Part 1 - Practice on DoS data from Census and HEIS	
Subject	Application of basic SAE methodologies from activity 2.1.1 on DoS data.	
	Direct estimators	
	 Indirect synthetic estimators 	
	Indirect composite estimators	
Methods	Peer-to-peer work	
Resources	MS, BC	
Duration	4 Days	
Output	Practice on basic smoothing on DoS data obtained taking outset in DoS	
	data for poverty statistics (using R)	
Time schedule	May 2023	

Activity 2.1.5	Applying theory and lessons learned on DoS data - Part 2 - Practice on DoS data from Census, HEIS, and administrative data		
Subject	 Application of advanced SAE modeling from activity 2.1.2 on DoS data. Modeling building Model diagnostics Sensitivity analysis Validation Quality assessment of results Comparisons with direct and other SAE estimates, such as those obtained in the World Bank project from 2017 		
Methods	Peer-to-peer work		
Resources	MS, BC		
Duration	4 Days		
Output	Practice on model-based small area estimations on DoS data obtained taking outset in DoS data for poverty statistics (using R)		
Time schedule	June 2023		

Bibliography

Elbers C., J.O. Lanjouw and P. Lanjouw, 2003. "Micro-level estimation of poverty and inequality," Econometrica 71(1): 355–364.

Céline Ferré and Matthew Wai Poi, 2020, Jordan Poverty Map, An Application of Small Area Estimation, METHODOLOGICAL REPORT.

Annex 1. Terms of Reference

EU Twinning Project JO 21 ENI ST 01 22

Component 2:

Reviewing and developing the source data and methodology for producing Small Area Statistics

Activity 2.0:

Assessment of current situation, defining of indicator and drafting the first rolling work plan

Dates: 17 - 20 October 2022

0. Objective and Mandatory results for the component

<u>Objective:</u>

To review and develop the source data and methodology for producing Small Area Statistics (SAS).

Mandatory results (MR):

- MR 2.1: Pilot project to assess inclusion of administrative and other external data sources in the development of SAS and action plan developed.
- **MR 2.2:** Develop methodology for producing SAS on poverty and provide recommendations on how this methodology can be applied to other areas within the DoS.
- MR 2.3: Develop training programs and manuals for use in the DoS and partner institutions based on pilot project outcomes.
- MR 2.4: Implement communication strategy with stakeholders on strengthening small area statistics.

1. Purpose of the activity

- Assessing the current situation for Small Area Statistics in Jordan
- Defining Indicators for each Mandatory Results (MR's)
- Drafting the first rolling work plan for developing the source data and methodology for producing Small Area Statistics

2. Expected output of the activity

Data and methodology for producing Small Area Statistics (SAS):

- The Twinning team, MS experts and DoS experts will obtain a common understanding of current situation, vision and challenges in relation to Small Area Statistics in Jordan.
- The level of compliance with relevant EU regulations and other international standards will be assessed.

Defining Indicators:

• Indicators will be defined for each MR.

In order to monitor the project implementation, indicators will be defined. The KPI's will follow the SMART principle (Specific, Measurable, Attainable, Relevant and Time bound). The KPI's will allow DoS, the Twinning Team and the Project Steering Committee (SC) to monitor and follow the progress in the project.

Drafting the first rolling work plan for the project:

- Headlines and brief descriptions to all planned activities Expect number of activities in component 2: 13 Short Terms Expert Mission; 1 Study visits
- Sequence and timing of activities for the first six month of the project
- Draft detailed description of activities for the first six months, in accordance with EU requirement using the template in table 1.

Activity number	Title
Subject	
Methods	
Resources	
Duration	
Output	
Time schedule	

Table 1: Twinning template for describing activities

3. Resources

Translation and interpretation will be provide throughout the activity

4. Overall agenda

Day 1: Current status in Jordan – needs and challenges related to the component

Day 2: Defining indicators for each MR

Day 3: Based on Indicator the first rolling work plan of activities in Jordan and study visits

Day 4: Summary and conclusions

5. Background information from the Fiche

Background and justification

Small Area Statistics (SAS) as a major policy requirement, particularly in measuring poverty rates and monitoring Sustainable Development Goals (SDGs). DoS surveys, particularly the Household Expenditure and Income Survey (HEIS) were not designed to provide reliable statistics at the small area levels required by policy makers. New methodologies, therefore, need to be introduced, including identifying new data sources and developing model-based techniques where necessary, to provide robust estimations with the required geographic breakdowns.

Components and results per component

Component 2: To review and develop the source data and methodology for producing Small Area Statistics

SAS, particularly in relation to poverty are derived from the HEIS, using weights from the CoP. The HEIS was designed at governate level and is not appropriate for developing statistics for smaller areas, as required by policy makers. The situation has been exacerbated by the influx of Syrian refugees who are included in the CoP totals, but where updated information is lacking since 2015, on their movements within the Kingdom and their current place of residence. Reliable representative samples, need to be developed for the purposes of compiling robust small area statistics. This component focuses on developing statistical methodologies for estimating small area statistics on poverty, and expanding the methodology to other statistical areas within the DoS. SAS data are a crucially important policy requirement for the purposes of monitoring Sustainable Development Indicators (SDIs).

Component 2 is divided into four sub-components with mandatory results applicable to each:

Mandatory Result 2.1: Undertake pilot project to investigate the inclusion of administrative and other external data sources in the development of small area statistics and develop an action plan

Developing SAS with the geographic breakdowns required by policy makers requires new data sources. Administrative data represents the obvious and most efficient solution as the collection of the requisite data by the DoS would be costly and could not be implemented within a short timeframe. However, developing comprehensive SAS can only take place over a long-time horizon – hence, a pilot project concentrating on key policy requirements is recommended. Poverty statistics and adherence to SDGs have been identified by users as priorities from a policy perspective. This sub-component will investigate available data sources, based on the inventory prepared under sub-component 1.1 above, assess their suitability, and detail the steps for inclusion in DoS estimates, where feasible. Key data sources for investigation include the Ministry of Social Development, while the CVDB (a non-government entity) is already compiling data at small area level. It may be necessary to recommend further data collection by the DoS, if robust estimates cannot be compiled using available administrative data.

> Mandatory Result 2.2: Develop methodology for producing SAS on poverty and recommend how this can be extended to other statistical domains

There is an increasing demand among policy makers for SAS. Often data at this level of disaggregation are not available from existing surveys as sample sizes are too small. This is especially true for Jordan in the area of poverty statistics. The key source of data are the HEIS, which was designed to produce statistics at governate level, and not at the smaller municipality

areas required by policy makers. Reliable representative samples, therefore, do not exist for the purposes of compiling robust small area statistics. The lack of reliable data is compounded by significant changes in the population, arising from political instability in the wider region. Better methodologies for estimating small area statistics are needed, which will, of necessity, require alternative data sources and may entail the use of modelling techniques. This sub-component assesses how administrative data can be incorporated in small area statistics and formulates an action plan. It will focus on enhancing poverty statistics (needed in measuring adherence to SDGs) which have been identified as a policy priority.

Developing reliable SAS usually requires the combination of a variety of data sources when survey data available to compilers are not designed for estimation at the level of disaggregation required. Administrative data can play a key role, but these data often need to be supplemented by model-based estimates. This sub-component will build on the work undertaken under 2.1 above. It will investigate the potential for developing a new methodology based on integrating available DoS and administrative data and augmenting these with statistical models, where necessary. The potential to extend any new methodologies adopted to other statistical domains will also be assessed— e.g., this may facilitate the development of spatial data.

> Mandatory Result 2.3: Develop training programmes and manuals for use within the DoS and with external counterparts

The development of SAS in the DoS will require extensive training, including the possible development of expertise, not currently available - e.g., in the area of modelling. Based on the results of sub-components 2.1 and 2.2 above, a comprehensive training plan needs to be developed including workshops, manuals and other training materials

> Mandatory Result 2.4: Implement communication strategy with stakeholders on strengthening small area statistics

SAS are a key policy requirement, particularly for government planning and monitoring compliance with SDGs. Given the importance of this information, the DoS role in providing the necessary data offers an excellent opportunity to enhance its public profile. This role needs to be highlighted and supported by all stakeholders at senior level. A comprehensive communications strategy, therefore, needs to be developed to inform users, media and the general public of the importance of DoS statistics in this regard.

Date	Place	Торіс
Monday – 17 Oct.	Meeting	Welcoming, acquaintance, program of the week
	room	(Mohammad Khalaf and Charlotte Nielsen)
10:00-12:00		BC: Introduction to component 2 – including introduction to
13:00-14:30		DoS and stakeholders
		BC: Overview of small area statistics in Jordan – status, needs and challenges
		Discussion, questions and remarks
Tuesday - 18 Oct.	Meeting	MS: Summary and conclusion from day 1
10:00-12:00	room	MS: Introduction to relevant EU best practice
13:00-14:30		Defining Indicators
Wednesday - 19 Oct.	Meeting room	MS: Summary and conclusion from day 2
10:00-12:00	100111	
13:00-14:30		Drafting the first rolling work plan
Thursday - 20 Oct.	Meeting	MS: Summary and conclusion from day 3
	room	Summarizing
10:00-12:00		Presentation of the outline for Indicators and working plan for
13:00-14:30		the management
		Final remarks and thanks

Annex 2: Programme for the mission

Abbreviations:

MS = *EU* Member State (Denmark, Germany, Italy, Lithuania, Finland); *DoS* = Department of Statistics, Jordan

Annex 3. Persons met

MS Short Term Experts (STE's):

PhD Andrius Čiginas, Specialist at Data Technology Development Group, Statistics Lithuania (MS Component Leader)

Mr. Čiginas is an expert in sample surveys and small area estimation for social and business statistics. He is strong in methodological and practical implementations of sample surveys and has vast experience in integrating administrative sources into surveys, and developing small area estimation modelling techniques. Mr. Čiginas has professional experience of more than 17 years in social and business sample surveys of official statistics, including methodological and hands-on experience with small area estimation in Statistics on Income and Living Conditions (SILC) and Labor Force surveys (LFS), and research based on SILC and LFS data. He has extensive experiences in delivering training.

Mr. Michele D'Alo, Head of a team in charge of small areas estimation and integration of data sources, The Italian National Institute of Statistics.

Mr. d'Alo has a long and established experience in the field of Small Area Estimation (SAE). Application of SAE methods and techniques to the main household surveys. Methodological issues related to multi-source data frameworks. Use of SAE to produce estimates of SDG indicators at very disaggregated level. Long and established experience in the field of Small Area Estimation (SAE). Application of SAE methods and techniques to the main household surveys. Methodological issues related to multi-source data frameworks.

DoS experts:

- 1. Ms. Fatima Awamreh, Head of Household Expenditure and Income Division (Component Leader) E-mail: <u>Fatmeh.Awamreh@DOS.GOV.JO</u>
- 2. Mr. Omar Hakouz, National Accounts expert DoS E-mail: <u>Omar.Hakouz@DOS.GOV.JO</u>
- 3. Mr. Mohammad Al-Jundi, Director of Household and Population Surveys Directory E-mail: Mohammad.AlJundi@DOS.GOV.JO
- 4. Mr. Mohammad Tawalbeh, Director of Agricultural and Environmental Statistics Directory E-mail: <u>Mohammad.Tawalbeh@DOS.GOV.JO</u>
- 5. Mr. Mohammad Abdel Razzaq, Director of Economic Surveys Directory E-mail: <u>Mohammad.AbdelRazzaq@DOS.GOV.JO</u>
- 6. Ms. Ghaida Khasawneh, Director of Methodologies and Studies Directory E-mail: <u>ha.Khasawneh@DOS.GOV.JO</u>
- 7. Mr. Safwat Al-Radaideh, Head of Administrative Data Department, poverty file E-mail: <u>Safwat.Radaideh@DOS.GOV.JO</u>
- 8. Mr. Mohammad Omari , Administrative Data Department E-mail: <u>Mohammad.Omari@DOS.GOV.JO</u>
- 9. Ms. Sana Al-Momani, Household and Population Surveys Directory E-mail: <u>Sana.AlMomani@DOS.GOV.JO</u>

Twinning team:

Eng. Mohammad Khalaf, Director of Sustainable Development Unit, Department of Statistics, Jordan (RTA Counterpart) E-mail: <u>Mohammed.Khalaf@DOS.GOV.JO</u> Ms. Charlotte Nielsen (RTA) E-mail: <u>cln@dst.dk</u> Ms. Zaina Amireh (Language Assistant) E-mail: <u>zainaamireh3@gmail.com</u>