TWINNING CONTRACT

BA 17 IPA ST 01 20

Further Support to the Reform of Statistics System in Bosnia and Herzegovina

MISSION REPORT

Activity 4D: Structure of earnings pilot survey
Component 4 – Labour Market Statistics

Mission carried out by
Jesper Moltrup-Nielsen, Statistics Denmark
Kirstine Andreasen, Statistics Denmark

06-09 March 2023

Version: Final

Funded by the European Union

FURTHER SUPPORT TO THE REFORM OF STATISTICS SYSTEM IN BIH

Expert contact information

Jesper Moltrup-Nielsen,
Statistics Denmark
Copenhagen, Denmark
Email: jmn@dst.dk

Kirstine Andreasen
Statistics Denmark
Copenhagen, Denmark
Email: kir@dst.dk

Table of contents

List of Abbreviations

BHAS	Agency for Statistics of Bosnia and Herzegovina
BiH	Bosnia and Herzegovina
CBBH	Central Bank of Bosnia and Herzegovina
EC	European Commission
EU	European Union
FBiH	Federation of Bosnia and Herzegovina
FIS	Institute for Statistics of Federation of Bosnia and Herzegovina
MS	EU Member State
RSIS	Institute for Statistics of Republika Srpska
RTA	Resident Twinning Adviser
ToR	Terms of Reference

1. General comments

This mission report was prepared within the EU Twinning Project "Further Support to the Reform of the Statistics System in Bosnia and Herzegovina". It was the fourth mission to be devoted to Structure of earnings pilot survey within Component 4 - Labour Market Statistics, of the Project.

The purposes of the mission were:

- Training of BC staff on determination of sample frame, calculation of all necessary indicators, including specific indicators (e.g. Variable 4.2) and also on determination in detail of editing and imputations;
- Finalising the analysis of data and work on outline structure of dissemination report.

The consultants would like to express their thanks to all officials and individuals from Bosnia and Herzegovina met for the kind support and valuable information, and which highly facilitated the work of the consultant.

This views and observations stated in this report are those of the consultant and do not necessarily correspond to the views of EU, BHAS, FIS, RSIS, CBBH, Statistics Denmark, Statistics Finland, Statistics Sweden or The Italian National Institute of Statistics.

2. Assessment and results

Presentation of results based on pilot data
During the mission, Vladimir Mijović presented data, indicators and results from the pilot conducted by the statistical institutions in BiH. Data from 38 enterprises and 1122 employees had been collected, validated and imputed. The results were as foreseen beforehand and therefore this is from the outset a good achievement on behalf of the BiH institutions.

A number of basic variables were computed to be able to calculate all indicators and to do breakdowns in different tables with a variety of classification variables. A number of graphs and tables was shown for indicators and basic variables.

After the completion of the pilot and the finalisation of the analysis, it is the common view of the participants of this mission, that the statistical institutions in BiH have the necessary skills to produce the needed indicators for delivering data on SES once a full-scale survey has been conducted.

Discussion on ISCO codes and descriptions
Results showed that many companies had found it difficult to fill in codes. The companies had an option to describe occupations in text as well as code. It seemed that it was easier to fill in the description. The codes was kept on two digit-level and very general. The conclusion was that the enterprise found it easier to type in the description. If they were giving the opportunity to give the 4-digit code, it might have been easier. However, the drawback of the 4-digit approach is that it gives a very long list of suggestions that could be confusing in its own right. The intension for the SES 2022 is to make the field for description mandatory.

Data processing
Data processing had been done in SPSS. Missing values and range of values were checked. Also consistency of variables was checked. One company did lack two variables for all employees. They were contacted, but they did not want to provide an answer. Therefore, values were imputed instead. A possible solution to this problem could be to make a form in the full SES where it is not possible to submit a response with some variables missing. Rules for validation from Eurostat were applied. Validation rules must be further developed when the next survey in full scale is conducted

Dissemination report
During the mission MS experts made a draft of a possible structure of a dissemination report, that could become the visible output of the pilot. This report was agreed upon during the first mission.

The draft included suggestions on possible graphs to use, based on the presentation made by the BiH statistical institutions on results from the pilot survey. The draft also had some preliminary textual descriptions that could be of interest to users, and also give the necessary information to readers about the limitations of the pilot study. There was a discussion about the risk that users will take the results for more final than the data can be taken to account for. This is the reason that it is important to describe the limitations clearly.

The draft report is shared with the project group for further work and finalisation after the mission is completed.

Sampling
Since it had been an important goal to get further into the understanding on sampling for the SES, there were some discussion about this topic. It was decided that it would be advisable to have an entire mission devoted to sampling, where dedicated sampling experts would have the opportunity to dive into the finer details.

On this mission the MS experts presented the basic concepts on sampling from a statisticians' point of view. The first topic were on how to work with target population, sample frame and the sample selection itself.

Based on a table of the overall structure of units in BiH, from a business register publication the MS experts presented possible ways to make sound strata that is sustainable over time. The experts presented a variety of things to consider when designing strata. For example how to work with the population, based on our knowledge about e.g. NACE codes, and to make decisions about where units can be representatives of each other, or where it might be advisable to make take-all stratification. There were also a discussion about the risks of making too many take-all strata, since it limits the options to recalibrate weights if non-response is experienced.

There were a discussion about whether the table used for the population is the correct one, but the participants agreed that it had suited the purpose of presenting things to consider when working with the population for stratification.

Machine learning model for ISCO code prediction
The data collection setup foreseen in the SES in BiH consists of free text descriptions of occupations as well as self-reported ISCO codes. Since we see difficulties on behalf of respondents to report correct codes, it is considered to make the free text descriptions mandatory. This will give the statisticians in the BiH institutions the task to ISCO-code many jobs.

Because of this, the MS experts presented an idea during the mission on how to make a model for predicting the most likely ISCO code as an inspiration. The model is based on a machine learning model that is trained in a Python environment. The model itself is used on a trial basis in Statistics Denmark on related statistics.

Since there are a few prerequisites that needs to be in place for the model to be transformed to BiH, this serves mostly as an inspiration for something to look into for the future.

Along with the development of the tool, it would be needed to decide on a process for implementing the outcome, and how to use it as a supporting tool that could limit the time for manual processing while still overviewing the general behavior of the model.

3. Conclusions and recommendations

In conclusion, the MS experts find that it has been a productive mission, which has provided the participants with the input needed to calculate the needed indicators in terms of being able to conduct a full-scale survey. In addition, the mission have concluded the analysis of the data and provided the BC with the outline structure for the dissemination report.

The experts recommend that BiH statistical institutions:

- Finish dissemination report
- Consider whether to use the dissemination report only as an internal document, or to publish it on their webpage to be able to show to future users and data suppliers
- Make an active decision on which variables are to be mandatory for data providers, e.g. ISCO description
- Complete the validation rules to be used for the next survey in full scale

Action	Deadline	Responsible person
Examine the possibility to conduct a dedicated mission on survey and sampling with dedicated experts	March 2023	RTA

Annex 1. Terms of Reference for the current mission

Terms of Reference

EU Twinning Project BA 17 IPA ST 01 20

Component 4 – Labour Market Statistics
06 March (starting at 10:00 hrs) – 09 March 2023 (08 and 09 March ending at 12:00 hrs.)

Venue: BHAS meeting room, Zelenih beretki 26
71 000, Sarajevo
Bosnia and Herzegovina

Activity 4.D: Structure of earnings pilot survey

1. Mandatory result

The component on Labour Market Statistics has the following objectives and Indicator / Relevant Milestones / Internal deadlines:

Component 4: Labour markets statistics: Structure of earnings pilot survey conducted and respective indicators produced		
Objective as in Twinning Fiche:	Indicator / Relevant Milestones / Internal deadlines	Status
4.1 Questionnaire and methodology for implementing Structure of earnings pilot survey prepared (including preparation of sample survey)	Questionnaire to be ready by end of 2021	
4.2 Developed application for entering data for the needs of BH institutions. The population and sample design defined, sample drawn.	February 2022	
4.3 Pilot survey conducted and data entered	Data expected to be collected during April 2022 and to be entered in June 2022	
4.4 Results of pilot survey analyzed	September 2022	

4.5 New indicators of the Labour market statistics based on the Structure of Earnings Pilot Survey such as: gross and hour wages by age, sex, occupation, duration of the working hours and type of contract services produced	December 2022	
4.6 Staff of the Labour market statistics trained for implementing full Structure of Earning Survey	February 2023/ End of project	

2. **Purpose of the activity:**
   - Training of BC staff on determination of sample frame, calculation of all necessary indicators, including specific indicators (e.g. Variable 4.2) and also on determination in detail of editing and imputations;
   - Finalising the analysis of data and work on outline structure of dissemination report.

3. **Expected output of the activity:**
   - Calculation of all necessary indicators prepared; editing and imputations methods prepared;
   - Collected data is finalized and ready for working on output products;
   - An outline structure for the final dissemination report is ready with all necessary indicators;

Annex 2. Persons met

MS experts:
Jesper Moltrup-Nielsen, Statistics Denmark
Kirstine Andreasen, Statistics Denmark

BHAS:
Radoslav Ćorović, Assistant Director in The Sector for Demography and Social Statistics
Vladimir Mijović, Head of Labour Market Department
Svjetlana Kezunović, Senior Officer in Labour Market Department
Tijana Popić, Senior Adviser in Labour Market Department

FIS:
Emina Šabanadžović, Senior Adviser for Labour Cost and Earning Statistics
Samka Avdić, Head of Labour Market Department
Edina Mehidić, Sampling Expert

RSIS:
Biljana Glušac, Senior Statistician for Labour, Wages and Employment Statistics

RTA Team:
Niels Madsen, RTA
Larisa Muslimovic, RTAA
Haris Imamovic, Interpreter

The content of this report is approved by: (sign and date)

MS Expert

MS Expert

BHAS Component leader
Svjetlana Kezunović

FIS Component leader
Emina Šabanadžović

RSIS Component leader
Biljana Glušac

RTA
Niels Madsen