**TWINNING PROJECT**

**Support to Development Process in the State Statistics Service of Ukraine with the Objective to Enhance its Capacity and Production**

**Ukraine**



**MISSION REPORT**

**on**

*Component 11: Capital and direct investment*

**Activity 11.4: Final mission**

Mission carried out by Andrius Čiginas and Tomas Rudys,

Statistics Lithuania

 13 - 16 October 2015

Version: Draft

|  |  |  |
| --- | --- | --- |
| Q:\PPT\SAMLING\JV\Logo2013\LogoUkPrint.png |  | logo_insee |
|  |  |  |

***IPA 2009***

*Author’s name, address, e-mail (keep the relevant information)*

*Andrius Čiginas*

*Tomas Rudys*

*Statistics Lithuania*

*29 Gedimino Avenue*

*LT-01500, Vilnius*

*Lithuania*

*Tel: (8 5) 236 4800*

*Email:* *Andrius.Ciginas@stat.gov.lt*; *Tomas.Rudys@stat.gov.lt*

**Table of contents**

[Executive summary 4](#_Toc434498877)

[1. General comments 4](#_Toc434498878)

[2. Assessment and results 4](#_Toc434498879)

 [2.1. Reliability of the sample survey results 5](#_Toc434498880)

 [2.2. Non-response and outliers 5](#_Toc434498881)

[2.3. Survey planning and estimation process in the future 5](#_Toc434498881)

[3. Conclusions and recommendations 6](#_Toc434498882)

[4. Final remarks 7](#_Toc434498882)

[Annex 1. Terms of Reference 8](#_Toc434498883)

[Annex 2. Mission programme 9](#_Toc434498883)

**List of Abbreviations**

ToR Terms of Reference

SSSU State Statistics Service of Ukraine

SL Statistics Lithuania

# Executive summary

The activities of component 11.4 will support SSSU in analysis of the estimation results. Best practices of data processing methods of all previous missions are summarized. In particular, the use of auxiliary information is emphasized at the estimation stage.

The overall purpose of the mission was:

* A comparison of sample survey estimates with that of previous (actual) methodology.
* Non-response and outliers treatment issues.
* Structure of sampling design and estimation stages; process coordination.

The activities in this mission included:

1. Discussion of the results presented by SSSU – comparison of the sample survey estimates with results based on previous methodology. Also, for a comparison, the corresponding sample survey results were presented by SL experts.
2. Several outliers treatment methods were practically tested.
3. Estimates for territories (additional estimation domains) were presented by SL experts.
4. Sampling design and estimation steps were overviewed once again. Some possible simplifications were discussed.

# 1. General comments

This mission report was prepared within the Twinning Project „Support to Development Process in the State Statistics Service of Ukraine with the Objective to enhance its Capacity and Production”. It was the last mission of SL experts to be devoted to the Capital investment within Component 11 of the project. The mission is aimed at preparing the sample data for estimation and inferences from them.

The concrete objectives of the mission were:

* share SL and EU practices in sample data quality: by means of non-response and outliers treatment;
* implement estimation strategies, which use auxiliary information, relevant for SSSU;
* summarize results of all previous missions.

The views and observations stated in this report are those of the experts and do not necessarily correspond to the views of EU, SSSU or SL.

# 2. Assessment and results

The assessment was carried out in accordance with the topics:

* 1. Reliability of the sample survey results.
	2. Non-response and outliers.
	3. Survey planning and estimation process in the future.

#

# 2.1. Reliability of the sample survey results

The experts of SSSU presented the comparison of the results obtained from sample survey and from previous methodology. For some of estimation domains (NACE activities), larger differences between the estimates were observed. In some cases, these differences are caused by outliers, which inflated estimates of the coefficients of variation. Corresponding estimates, presented by experts of SL, use auxiliary information (data from annual survey) for some of the estimation domains. Therefore, these estimates appeared more consistent with the previous methodology results. However, it was concluded that more work should be done with sample data quality – non-response and outliers, in order to avoid instability of estimates.

The estimates for territories were also calculated and compared. Their reliability appeared almost independent of the sampling design (which was based on the stratification by NACE activities), but influence of outliers is observed too.

# 2.2. Non-response and outliers

The Monte-Carlo study of experts of SL showed that, in cases of large non-response rates (especially in the first quarter of the survey), the Horvitz-Thompson estimators imply large biases of the results. In this situation, the special cases of calibration estimators – ratio and regression estimators – reduce the impact of non-response. Single sample results agree with this conclusion.

First, several simple rules were practically tested in order to detect outliers and neutralize their impact. For instance, simple quartiles rules appeared strongly dependent on estimation domains, thus cannot be applied automatically. But more sophisticated Winsorization rules were very efficient. The estimates were significantly improved.

**2.3. Survey planning and estimation process in the future**

The main steps of the sample survey were overviewed and possible problems of their implementation were discussed. It was noted by experts of SSSU that there is the need to have estimates at detailed level of 3 NACE digits for some of economical sectors. The experts of SL pointed also that here the work coordination between the subject matter specialists and surveys sampling specialists is important. That is, a list of technical tasks should be prepared by subject matter specialists concerning estimation aims of the survey.

It was agreed that further practice with calculation of estimates (of the third and fourth quarter) are necessarily for SSSU before launching the sample survey alone.

1. **Conclusions and recommendations**

We present the final conclusions in the form of sampling design and estimation steps.

Planning of the sample:

1. Formation of the target population. It should be decided which of small enterprises can be excluded from the survey (due to their insignificancy). Here participation of the survey sampling specialists is preferable.
2. Estimation domains must be clearly indicated.
3. Available quarterly micro-data on investment of the current year and the corresponding annual data of the previous year are linked to the current list of the target population. Using information on states of enterprises, zero values for the investment variables can be imputed for some of enterprises.
4. Two auxiliary variables are formed: $v$ – the historical annual investment; $z$ – the mean of available quarterly investment data (two or three current quarters). Then, by the rule of quartiles, the outlying enterprises are detected by each of the variables, and separate “take all” strata are formed.
5. One more strata of “take all” enterprises are formed from enterprises with large number of employees.
6. Singe auxiliary variable $x$ is defined using $z$ and $v$: set $x=z$, and then, if $x$ is empty then put $x=v/4$.
7. The enterprises with empty $x$ form the separate strata in each of the estimation domains (by NACE activity). The sample sizes for these strata are assigned later.
8. Small estimation domains form separate “take all” strata.
9. A part of the available sample size is used already for the “take all” strata. The remaining sample size is iteratively allocated to estimation domains with the aim to ensure similar accuracy of the future estimates (use the variable $x$).
10. In each estimation domain, the number of strata can be obtained by artificial simulations. For simplicity, these numbers can be taken from previous analysis. Similarly, using “power-stratification” for determination of strata bounds, the results of the previous year can be used.
11. In each of the estimation domains, the structure of strata is defined. Here Neyman allocation is performed using the auxiliary variable $x$.
12. Proportionally to the sample size applied for the estimation domains, the sample sizes are assigned for the strata with enterprises with unknown $x$.
13. The sampling design is constructed, and the simple random sample is drawn in each of the stratum.

 Estimation:

1. The auxiliary information variable is prepared: $x$ is the sum of four quarters of the previous year survey. Where $x$ is empty, the corresponding data on investment of the annual survey of the previous year is used.
2. For the survey variables (sample of the quarter), zero values are imputed for some of non-respondents depending on their known states of non-response.
3. Expert imputation should be performed by subject matter specialists in cases of influential non-respondents (enterprises with large expected investment). Reweighting is applied for the remaining non-respondents.
4. To detect outliers, and thus possible errors in the data, the list of suspicious sample data is prepared (using, for instance, the quartiles rules) by the survey sampling specialists and given to the subject matter specialists for analysis. At this point, a cooperation of these specialists is the most important.
5. When the survey data are assumed correct, Winsorization technique is applied to reduce an impact of outliers with larger than the unit weights.
6. Ratio estimators are applied to the estimation domains where the survey variables and the auxiliary information $x$ are sufficiently correlated. In these cases, non-response biases are reduced. For other estimation domains, the direct Horvitz-Thomson estimators are applied.
7. If the resulting estimates appear inconsistent from macro-economic considerations, the subject matter and survey sampling specialists might return to the micro-level analysis.

## Final remarks

1. During the project, efficient cooperation between experts of SSSU and SL was achieved. In that follows, the sampling design was effectively modelled thanks to the available auxiliary information. Further, the pilot calculations showed that SSSU has no reasons not to launch the quarterly sample survey.
2. SSSU should pay more attention to the staff of survey statisticians. For instance, for the quarterly survey on Investment, considering a complexity of the sampling design, minimum 2 weeks are necessarily for the sample selection, and minimum 1 week for calculation of estimates for each quarter.
3. Experts from SL are ready to cooperate in the future. In particular, in preparation of methodologies or even for technical solutions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Q:\PPT\SAMLING\JV\Logo2013\LogoUkPrint.png |  |  | statis |  |

**European Union Twinning Project**

**Support to Development Process in the State Statistics Service of Ukraine with the Objective to Enhance its Capacity and Production**

**Twinning No.: UA/13/ENP/ST/38**

**Terms of Reference**

for a short-term mission to the State Statistics Service of Ukraine

13 – 15 October 2015

*Component 11 Capital and direct investment Activity 11.4 Final mission.*

**General information**

Statistics Denmark in partnership with INSEE France, Statistics Lithuania, Statistics Finland, Central Statistical Bureau of Latvia, is leading the EU-Twinning project on “Support to Development Process in the State Statistics Service of Ukraine with the Objective to Enhance its Capacity and Production” in Ukraine. The beneficiary is the State Statistics Service of Ukraine.

This activity is implemented under the Component 11 Capital and direct investment.

**Purpose of the mission**

Present the estimations, resulting from the capital investment survey within the designed sample of small enterprises for the 1st and 2nd quarters. (SSSU)

Analyze the results within the designed sample of small enterprises for the 1st and 2nd quarters to conduct an experimental capital investment survey covering the 2nd quarter of 2015.

Discussion and expert recommendations on processing of “no answers”.

Get expert recommendations on designing a sample of small enterprises, incl. by institutional sectors of economy S.12 Financial corporations and S.13 Sector of General public governance.

Get expert recommendations on methods for evaluating results; accuracy and quality of such evaluations.

**Consultant and partner:**

Mr Andrius Čiginas (Statistics Lithuania)

Mr Tomas Rudys (Statistics Lithuania)

**Beneficiary Institution:**

I.Petrenko, Acting Director, Production Statistics Department, SSSU;

O.Myslinsky, Head of Division for Statistics of Capital investments and fixed assets, Department for Statistical production, SSSU

O.Khmelyovska, Chief -economist at the Division for Statistics of Capital investments and fixed assets, Department for Statistical production, SSSU

M.Gulenko, Chief -economist at the Division for Statistics of Capital investments and fixed assets, Department for Statistical production, SSSU

A.Tovchenko, Head of Division of mathematic processing methods and analysis of statistical data, Department for Statistical infrastructure, SSSU

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Описание: Q:\PPT\SAMLING\JV\Logo2013\LogoUkPrint.png |  |  | Описание: statis |  |

**European Union Twinning Project**

**Support to Development Process in the State Statistics Service of Ukraine with the Objective to Enhance its Capacity and Production**

**Twinning No.: UA/13/ENP/ST/38**

**MISSION PROGRAMME**

Working meetings between the EU experts and the experts of the Production Statistics Department & Department for Statistical infrastructure (SSSU) within the framework of the Twinning Project “Support to Development Process in the State Statistics Service of Ukraine with the Objective to Enhance its Capacity and Production”

13-16 October 2015

**Component 11 Capital and direct investment Activity 11.4 Working mission.**

**Experts:**

**Mr Andrius Čiginas, Statistics Lithuania**

**Mr Tomas Rudys, Statistics Lithuania**

**Date: 13-16 October 2015**

**Working language: Russian**

|  |  |  |
| --- | --- | --- |
|  | **Morning** | **Afternoon** |
| Tuesday13/10/2015 | **10:00 Arrival to the SSSU** At the entrance to the SSSU you will be picked up by Kateryna Zhulay, Protocol Department**10:00 – 13:00**SSSU: Presentation of estimations, resulting from the capital investment survey within the designed sample of small enterprises for the 1st and 2nd quarters.Analyzing the results within the designed sample of small enterprises for the 1st and 2nd quarters to conduct an experimental capital investment survey covering the 2nd quarter of 2015. | 14:30 – 16:30 To be continued |
| Wednesday14/10/2015 | **10:00 Arrival to the SSSU** At the entrance to the SSSU you will be picked up by Kateryna Zhulay, Protocol Department**10:00 – 13:00**Discussion and expert recommendations on processing of “no answers”. Expert recommendations on designing a sample of small enterprises, incl. by institutional sectors of economy S.12 Financial corporations and S.13 Sector of General public governance.  | 14:30 – 16:30 To be continued |
| Thursday15/10/2015 | **10:00 Arrival to the SSSU** At the entrance to the SSSU you will be picked up by Kateryna Zhulay, Protocol Department**10:00 – 13:00**Expert recommendations on methods for evaluating results; accuracy and quality of such evaluations. Discussion, debriefing | 14:30 – 16:30 To be continued |
| Friday16/10/2015 | **10:00 Arrival to the SSSU** At the entrance to the SSSU you will be picked up by Kateryna Zhulay, Protocol Department**xx:00 – 13:00** | 14:30 – 16:30 To be continued |

**Participants**

Mr Andrius Čiginas, Statistics Lithuania

Mr Tomas Rudys, Statistics Lithuania

Ms I.Petrenko, Acting Director, Production Statistics Department, SSSU;

Mr O.Myslinsky, Head of Division for Statistics of Capital investments and fixed assets, Department for Statistical production, SSSU

O.Khmelyovska, Chief -economist at the Division for Statistics of Capital investments and fixed assets, Department for Statistical production, SSSU

M.Gulenko, Chief -economist at the Division for Statistics of Capital investments and fixed assets, Department for Statistical production, SSSU

A.Tovchenko, Head of Division of mathematic processing methods and analysis of statistical data, Department for Statistical infrastructure, SSSU