**Title of Paper: Quality monitoring and improvement of innovation survey in SSO of Republic of Macedonia**

Tatjana Drangovska M.Sc., State Statistical Office of the Republic of Macedonia,

tatjana.drangovska@stat.gov.mk

**Abstract**

*Innovation has a central role in the process of output creation and productivity growth. Data on innovation are widely used by policy makers for policy creation and by researchers for estimating different aspects of the innovation process. Therefore, production of high-quality data on innovation is of particular interest. From 2013, the SSO started conducting a survey on innovation. In order to ensure international comparability, the standard CIS questionnaire and Oslo Manual are used. For quality monitoring of the outputs of innovation data and for making the figures transparent and easily accessible to the users, ESQRS, DESAP self-assessment and national Quality Report are prepared. The purpose of the paper is to present the improvement of the innovation data through the period based on the key indicators calculated in quality reports used as input for the future work. At the same time, the paper will briefly present the methods undertaken for improving the quality of output data on innovation, for example method of cross-checking with data from other statistical surveys and the usage of data from administrative sources.*

**Keywords:** Innovation survey, quality reports, quality improvements.

**1. The need for collecting data on innovation**

It is widely accepted that innovation is a key driver in the process of creation of output and economic growth; this raises the need for production of data and indicators of innovation. In recent years, innovation surveys are a widely used source for innovation indicators are. They provide qualitative and quantitative data on innovation and on different aspects of the innovation process, such as: input in innovation, R&D activities and other activities different from R&D, innovation output and effect of introduction of innovations. Data obtained by the surveys are extensively taken into consideration by statisticians and policy makers to monitor and benchmark innovation performance, as well as by economists and researchers to explore and analyse the effects of innovation introduction and other related topics (Mairesse and Mohnen, 2010). Data obtained by the innovation survey in our country are used by policy makers for preparation of the national Strategy for Innovation and Strategy for Smart Specialisation, as well as for calculation of the Innovation Index. Other data users are academic researchers; their interest is mostly focused on studding the innovation process. International organizations are also one of the key users. All the above-mentioned different type of users raise the need for production of high-quality innovation data and indicators and systematic monitoring of the quality as a final phase of the statistical operation process.

**2. Innovation statistics in the SSO – background**

The SSO conducted its first pilot survey on innovation in 2005, for the reference period 2002-2004. The first regular survey was conducted in 2013 for the reference period 2010-2012. The questionnaire was first tested in the field and after that the regular survey was conducted in the same year. The results from the test showed that it is quite difficult for the reporting units to identify what is innovation for them, to make differentiation between different types of innovation and, in particular, the test showed problems with reporting of innovation expenditures. Based on the weaknesses identified with the test, improvements were made to the survey instruments. The standard form of the CIS[[1]](#footnote-1) questionnaire was used as an instrument for data collection and all definitions were in line with the Oslo Manual. The regular survey was conducted on a stratified sample; the stratification was done by number of employees and economic activity of business entities, mentioned as mandatory and voluntary in Commission Regulation (EU) No. 995/2012.[[2]](#footnote-2) The questionnaires were sent by post together with methodological explanations and samples of different types of innovations. On the SSO web site there were instructions for completing the questionnaire. The response rate was only 55.8%. The obtained results showed that for most of the innovation active enterprises it is hard to report the innovation expenditures and turnover from product innovations. Over 60% of the innovation active enterprises did not report any type of expenditure. The results were published on the SSO web site and were not delivered to Eurostat. A quality report was not prepared, only some quality indicators were calculated, and quality evaluation was done based on the indicators.

3. Preparation of Quality Reports and systematic monitoring of quality of innovation statistics in the SSO

As it is important to produce high-quality data, the monitoring of data quality is an issue of particular importance. For monitoring the quality of innovation data, quality reports are prepared, with the aim to assess the data characteristics by quality components and to present the data features according to the standard quality requirements. Quality reports, besides assuring quality assessment, also represent a starting point for quality improvements. For the reference period 2012-2014, three kinds of Quality Reports for Innovation survey were prepared: National Quality Report ESMS, ESQRS and DESAP self-assessment questionnaire. For the reference period 2014-2016, the reports are in the phase of preparation and only the key quality indicators are calculated.

*3.1. National Quality Report – Euro SDMX Metadata Standards (ESMS)*

The National Quality Report prepared for the Innovation Survey follows the Euro-SDMX Metadata Structure (ESMS) and consists of 21 concepts with a limited breakdown of sub-concepts. The Report contains information on methodology, quality and on the statistical production process in general of the innovation survey, and it is available on the SSO web site. Because it is a user-oriented report, the main aim is to inform the users of innovation data on quality aspects of the data and to give some general information regarding the innovation data. The accent of the Report is on quality assessment of survey outputs.

*3.2. ESS Standards for Quality Reporting (ESQRS)*

The monitoring and assessment of all quality components of the statistical production process is of high interest for survey managers in order for future steps for quality improvements of the data to be identified. ESQRS for the innovation survey was prepared with the aim to assess the data production process. The Report was delivered to Eurostat via eDamis. The Report consists of 12 concepts. Because the innovation survey is a sample survey, generally, the quality of innovation data is monitored through the non-response and imputation rate, as they reflect the reliability and accuracy of the data. For this reason, the main findings and the assessment of those concepts are presented in the paper. For monitoring of the non-response rate, un-weighted and weighted rates were calculated. The un-weighted rate was calculated in order to have a clear picture of how representative the survey results are and if a non-response survey should be undertaken. The weighted rate is used as an indicator for mapping of potential bias caused by non-response prior to any calibration method. The un-weighted rate, referring to the data of the sample, for the reference period 2012-2014, is 26%. The weighted non-response rate, which represents how well the data collection worked considering the population, is not significantly higher, 28.59%, and is not exceeding 30%. Following Eurostat’s recommendations, a non-response survey is not carried out, because the non-response rate does not exceed 30%. For deeper analyses and to have a clearer picture for this key indicator, the non-response rates are calculated by NACE classes. Analysed by NACE classes, the non-response rate is higher for services, compared to industry. The situation is the same for un-weighted and weighted rates.

**Table 1. Unit non-response rate (%), for innovation survey, by NACE classes, reference period 2012-2014**

|  |  |
| --- | --- |
|   | **2012-2014** |
| **Number of eligible units with no response - population** | **Total number of eligible units in the sample** | **Un-weighted unit non-response rate** | **Weighted unit non-response rate** |
| **Core NACE (B-C-D-E-46-H-J-K-71-72-73)** | 857 | 1500 | 26 | 28.59 |
| **Core industry (B\_C\_D\_E - excluding construction)** | 429 | 888 | 22.75 | 26.27 |
| **Core Services (46-H-J-K-71-72-73)** | 428 | 612 | 30.72 | 31.36 |

Source: State Statistical Office of Republic of Macedonia, ESQRS for innovation survey

Because of the impact of non-repose on the output quality, the main focus was put on improving the response rate. The calculated non-response rates for the last reference period are presented in Table 2. Compared to the previous reference period, for the last reference period 2014-2016, the non-response rate is significantly lower. The un-weighted and weighted rates are almost the same, 17%. The indicator shows the improvement of the data quality compared to the previous reference period. Improvement of response is ensured in industry as well as in services. Higher improvement regarding the response is achieved among the enterprises in service classes. The un-weighted non-response rate for service classes is reduced from 30% to 11%.

**Table 2. Unit non-response rate (%), for innovation survey, by NACE classes, reference period 2014-2016**

|  |  |
| --- | --- |
|   | **2014-2016** |
| **Number of eligible units with no response - population** | **Total number of eligible units in the sample** | **Un-weighted unit non-response rate** | **Weighted unit non-response rate** |
| **Core NACE (B-C-D-E-46-H-J-K-71-72-73)** | 542 | 1515 | 17.23 | 17.36 |
| **Core industry (B\_C\_D\_E - excluding construction)** | 358 | 870 | 21.49 | 22.03 |
| **Core Services (46-H-J-K-71-72-73)** | 184 | 645 | 11.47 | 12.28 |

Source: State Statistical Office of Republic of Macedonia, ESQRS for innovation survey

The non-response rates are calculated by size classes of enterprises for two reference periods. The rates calculated are presented in Figure 1. Comparing the figures from the two reference periods, the non-response is reduced for all three classes of enterprises, but the most significant improvement is achieved among small enterprises.

**Figure 1. Unit non-response rate (%), for innovation survey, by size classes, reference periods 2012-2014 and 2014-2016**



 Source: Calculation of author

Data imputation is done for missing, invalid and inconsistent data. The imputation was done after every attempt to get the information from the reporting unit. Only the un-weighted imputation rate for the key variables is calculated. The imputation rates for three metric survey indicators (total turnover in the last year of the reference period (t), share of the turnover due to new or improved products in the total turnover for product innovative enterprises and expenditures in intramural R&D) are shown in Table 3. For reference period 2012-2014, the imputation rate was not calculated for total turnover, because the question on turnover is omitted and data are taken from the statistical business register. The calculation of the imputation rates shows the reduction of the imputation rates for all key variables between the two reference periods.

**Table 3. Un-weighted imputation rates (%) for three metric variables of innovation survey, reference periods 2012-2014 and 2014-2016**

|  |  |  |
| --- | --- | --- |
| **Un-weighted imputation rate**  | **2012-2014** | **2014-2016** |
| **(TURN t/TOTAL)[[3]](#footnote-3)** | N/A | 2.0 |
| **(NEWMAR\_TURN/INPDT)[[4]](#footnote-4)** | 7.5 | 6.1 |
| **(RRDINX/INNOACT)[[5]](#footnote-5)** | 5.9 | 4.8 |

Source: State Statistical Office of Republic of Macedonia, ESQRS for innovation survey

As it is important to produce comparable data, the data comparability and accessibility are evaluated. Regarding the comparability concept, the innovation survey produces highly comparable data. For ensuring international comparability as mentioned before, the standard CIS harmonised questionnaire is used, only the question referring to world innovation is omitted. Regarding the coverage of NACE sectors and size classes, included were classes and sections corresponding to the Eurostat recommendations. Innovation data are available at aggregated level and, with special approval, as microdata sets. Data are disseminated on the SSO web site in News Release, different Publications and in the Database. Data are free of charge. Together with the data, there are methodological notes in order to help the users in interpretation of the indicators published. The data are available for researchers at micro level with special approval. In recent years, data gathered through the innovation survey are used in a lot of academic publications.

*3.3. DESAP self-assessment questionnaire*

Another report prepared for assessing the quality of the innovation data is DESAP self-assessment questionnaire. The DESAP was filled by staff that participate in different phases of the statistical production process.

With the DESAP questionnaires, the sample design process as a starting point for ensuring quality of all further processes was assessed with high quality. Regarding the issues of overcoverage, undercoverage and misclassification of the statistical units in the sample, no significant problems were identified. The careful selection and design of the survey concepts assured high quality regarding the timeliness, punctuality and comparability of the data. The DESAP questionnaire assessed high quality and fulfilment of all requirements regarding the timeliness and punctuality. The data are published following the EU requirements and according to the planned publication date. The data obtained by the innovation survey are assessed as highly comparable between non-geographical areas and over time.

On the other hand, the DESAP questionnaire identified some weaknesses regarding the production process. One of the weaknesses is the lack of available information on user satisfaction of key users of innovation data. There is available some general information on user satisfaction at the level of all data produced by the SSO, but not for different areas and surveys. Another weakness identified refers to the existence of errors in raw data. No validation program that performs plausibility checks is implemented in the collection process, and the data editing is done manually. There is also place for improvement of the non-response rate, which exceeds 15% and this has impact on the survey results. High item non-response rate was calculated for production of product innovation and innovation expenditures as one of key variables.

The data coherence is not assessed, because the survey is not carried out with different frequencies at the same time. The results from the DESAP are presented in Figure 1.

**Figure 1: DESAP self-assessment diagram for innovation survey, reference period 2012-2014**



Source: DESAP self-assessment diagram

**4. Steps undertaken to increase the data quality**

Based on the results and weaknesses identified during quality monitoring and based on the quality indicators calculated, some steps were undertaken for quality improvement of innovation data.

*4.1. Establishment of List of innovative enterprises*

For improvement of the sample design, appropriate identification of innovative enterprises and of the type of innovations introduced by the enterprises, a List of innovative enterprises is established. The source used for preparation of the List is the web site: [www.inovativnost.mk](http://www.inovativnost.mk). On the web site, there is information on: innovative enterprises in the country, detailed explanation about the types of innovations introduced, data on cooperation for development of innovation, data on investment made for implementation of innovation and, for some enterprises, information is available on funding source of the innovations. This List is regularly updated. The web sites of the enterprises are also used as a data source for updating the List. Most of the large enterprises and most innovative ones have web sites with available information about the introduction of innovation (mostly for product or process innovation).

*4.2. Cooperation with the Fund for Innovation and Technology Development*

For data checking and reducing non-response of data for receiving public support, cooperation has been established with the national Fund for Innovation and Technology Development. The SSO receives on a yearly basis a List with enterprises that received national financial support for introduction of innovation and information about the type of innovation for which the enterprise received financial support. The Fund is a national body established for offering financial support for introduction of innovation. Regarding the financial support received by the Horizon 2020 Programme for Research and Innovation, for mapping of enterprises which are part of the Programme and received financial support, the Programme web site is used. The web site contains information about the type of the Project and cooperation partners.

*4.3. Comparison of the results with the national R&D survey*

For reduction of item non-response on R&D expenditures and providing more reliable data, the survey results are compared with the results from the national R&D survey. The national survey is not used only for checking of R&D expenditures, but also for improving the sample.

**Future plans**

In future the SSO will continue with regular implementation of the survey on innovation. With the aim of ensuring quality through the whole production process of innovation statistics, the SSO will continue with preparation of quality reports, calculation of quality indicators and quality assessment of input, throughput and output of the innovation survey. Quality monitoring process is not performed only to assess the current data quality, but also to undertake future steps for quality improvement. For unit non-response rate, as one of the key indicators, also analysed in the paper, more effort will be put on ensuring high response rate. Most of the countries started using electronic questionnaire for data collection on innovation and the results showed improvement in data quality, and the SSO will make efforts to implement an electronic data collection questionnaire to improve the quality through the whole statistical production process. But mainly the aim of the implementation of an electronic questionnaire will be to improve the collection process and data processing, because a validation program that performs plausibility checks during the data collection will be implemented in the questionnaire in order to avoid manual data editing. Following the experience from other countries, for the first data collection process, maybe a mixed mode approach will be applied, i.e. a combination of online survey and postal mail. Because the innovation survey collects a wider range of data on different aspects of the innovation process, the most important data used for analysis of the input and output side of the innovation are the innovation expenditures and the turnover from innovation. The quality improvement of those data is planned to be done by data mining on web sites of companies and comparison with the financial statements of the companies.

**5. References**

ESTAT/D4/LA D (2014), ESS Guidelines for the implementation of the ESS priority quality indicators, Luxemburg. Available at: <http://ec.europa.eu/eurostat/documents/64157/4373903/02-ESS-Quality-and-performance-Indicators-2014.pdf/5c996003-b770-4a7c-9c2f-bf733e6b1f31>

ESTAT/D4 (2015), Single Integrated Metadata Structure v 2.0 – ESMS 2.0 and ESQRS 2.0, Luxemburg. Available at: <http://ec.europa.eu/eurostat/documents/64157/4373903/SIMS-2-0-Revised-standards-November-2015-ESSC-final.pdf/47c0b80d-0e19-4777-8f9e-28f89f82ce18>

Mairesse J. and Mohnen P. (2010), Using innovations surveys for econometric analysis, National Bureau of Economic Research Cambridge, Working paper 15857. Available at: <http://www.nber.org/papers/w15857>

1. Community Innovation Survey, common questionnaire developed by Eurostat and OECD. The questionnaire is used by EU and non-EU countries for data collection in order to ensure international comparability regarding innovation performance among business section, between countries. [↑](#footnote-ref-1)
2. Commission Regulation No.995/2012 concerning the production and development of Community statistics on science and technology. [↑](#footnote-ref-2)
3. Total turnover in the last year of the reference period (t) [↑](#footnote-ref-3)
4. Share of the turnover due to new or improved products in the total turnover for product innovative enterprises [↑](#footnote-ref-4)
5. Expenditures in intramural R&D [↑](#footnote-ref-5)