

26th Meeting of the Wiesbaden Group on Business Registers

Neuchâtel, 24 – 27 September 2018

Agne Bikauskaite, Zsolt Völfinger (Eurostat)

Session 8 - Output of Statistical Business Registers

Multinational enterprise groups in the EU – Dissemination from the EuroGroups Register

1. Introduction

Multinational enterprise groups play a major role in the EU economy. They contribute substantially in all EU countries to the production of goods and services, employment and investment.

Despite growing interest in the structure of multinational enterprise groups, Eurostat does not yet publish any European official statistics on that topic. To meet this information need, Eurostat started publishing experimental statistics on multinational enterprise groups in the EU and EFTA countries.

The statistics compiled are based on data from the EuroGroups Register (EGR) and aggregated by different breakdowns. The EGR is a statistical register of Eurostat and national statistical institutes (NSIs) of the EU and EFTA countries. The EGR covers the population and units of multinational enterprise groups present in the EU and EFTA. If one legal unit of a multinational group is in the EU or EFTA, the group is already in the scope of the EGR.

2. EGR production process and results

The EGR produces annual data, reflecting the status of the registered groups as of 31st December of the reference year. The collection and validation of the EGR data are complex and long processes. The EGR final data for the reference years are available for users 15 months after the end of each reference year.

To create the EGR, Eurostat collects input information on enterprise groups from the national statistical business registers of 28 EU and 4 EFTA countries, and from one commercial data source. National registers deliver micro data on the constituent units of the groups and on their relationships. The EGR also acquires data from one commercial data provider; this acquisition ensures coverage for units outside the EU+EFTA and consistency of the largest enterprise groups.

After validation of input data from the different sources the EGR calculates the consolidated picture on the legal units (companies), on their relationships and on their enterprises (the statistical representation of the legal units). Complementary and conflicting information from the sources are treated in the EGR process with predefined priority settings and business rules.

Having final information on legal units and relationships for a reference year the EGR calculates the enterprise group structures. The enterprise groups are built on those legal units, which are linked together by control relationships, where the voting rights are above 50%.

The final EGR data with harmonised picture on the groups are accessible only for statisticians producing national statistics in EU and EFTA NSIs and Central Banks.

The EGR process was redesigned to EGR 2.0 from 2012, the new process was fully implemented in 2015. The switch to EGR 2.0 improved both the EGR process and the IT system. Data for reference years 2014, 2015 and 2016 were produced completely with the EGR 2.0 system.

EGR 2.0 aims to cover all relevant multinational enterprise groups active in the EU. For reference year 2014 EGR produced data of 61 000 multinational enterprise groups, for 2015 data of 80 000 groups. For the last available reference year, 2016 EGR produced data on 111 000 groups covering 779 000 enterprises and 944 000 legal units.

EGR data enable information to be produced on the overall structure of multinational enterprise groups in the EU, while preserving the confidentiality of micro-data.

3. Experimental statistics in Eurostat

Eurostat started to publish experimental statistics in 2017. These statistics use new data sources and methods in an effort to expand and improve the way to respond to our users' needs that are not met yet by official statistics.

It should be noted that these statistics have not reached full maturity in terms of harmonisation, coverage or methodology. This is why they are always marked with a clearly visible experimental logo and accompanied by detailed methodological notes.

Users are reminded that experimental statistics are test data which should be used with caution. They do not fully comply with the standards of the robust legal and quality framework of European statistics.

The main purpose is to experiment and to seek user feedback. Eurostat will use the feedback to further improve the robustness of experimental statistics.

Users are invited to carefully read all explanations on the data sources and methodology used. There is always a risk with such publications, that users simply accept the data published on the website of Eurostat as European official statistics and do not read up carefully the attached disclaimers and methodological notes.

Eurostat website has a specific section for experimental statistics; this site covers all experimental publications disseminated up to now: <http://ec.europa.eu/eurostat/web/experimental-statistics>

4. Why are the EGR data published as experimental statistics?

The main reason to publish statistics on multinational enterprise groups from EGR as experimental statistics is the quality of the EGR database, in particular as regards the relative completeness and accuracy of its data.

The EGR has very broad coverage of large and medium-sized groups. However, its estimated coverage of small groups and their constituent enterprises stands at only 60%. This means it is not yet possible to reach very precise conclusions on the total population of multinational enterprise groups in the EU using EGR data.

The EGR is only one among various sources of national statistics on enterprise groups and foreign affiliates. EGR data are not always comparable with similar statistics on enterprise groups and affiliates in national publications. This is because national statistical authorities use additional information and procedures to produce national data.

Another reason for the experimental status of EGR data is the way data on group structures are produced. The EGR is a platform that merges the data of EU and EFTA countries with commercial

data. The final view on groups is established on the basis of sources' preference rules and the business rules that govern the EGR system. The outcome may thus differ from the national view on groups.

Combining multiple sources can also lead to errors in group structures. Despite all the validation work invested in the process, accuracy problems may arise in the EGR output. Some group structures may be broken or fragmented. Some input data, e.g. employment figures of enterprises can be missing.

Eurostat and the participating national statistical authorities are continuously working on the improvement of the EGR quality.

5. Dissemination article from EGR

Eurostat published one statistical article on multinational enterprise groups as experimental statistics. This article first was published with 2015 data in January 2018 and was replaced with 2016 data in July 2018. The article on the groups can be found under the Eurostat Statistics Explained website: http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Structure_of_multinational_enterprise_groups_in_the_EU

With the dissemination from EGR Eurostat would like to avoid producing overlapping data with European statistics on Foreign Affiliates (FATS). FATS are producing statistics on foreign controlled enterprises and on foreign affiliates:

- Foreign control of enterprises by controlling countries, by economic activity
- Foreign affiliates of EU enterprises and their characteristics

Eurostat coordinates closely the EGR output with the Eurostat publications on FATS. The EGR output tables were designed to be complementary to the FATS data tables, mostly focusing on the structure of the multinational enterprise groups in the EU and on the characteristics of the groups. Data on the statistical unit enterprise group are not published in FATS by group characteristics.

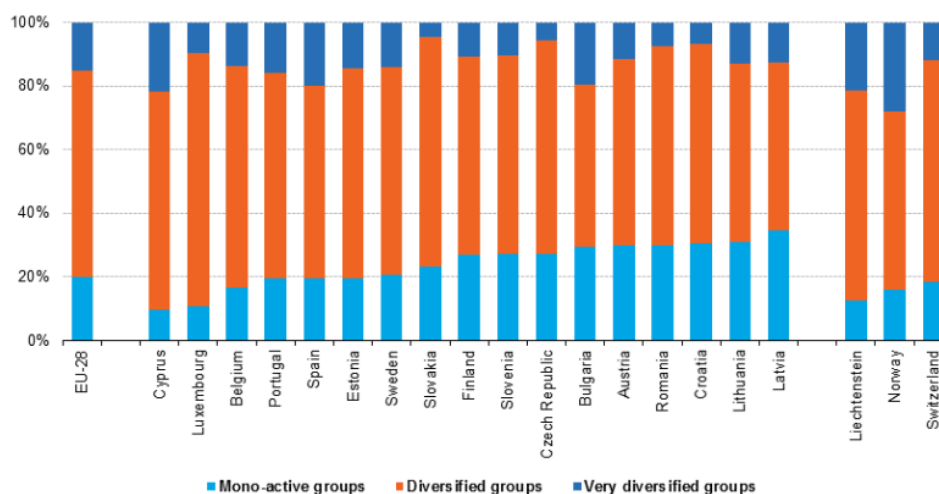
The EGR experimental statistics data tables and graphs are based on the following characteristics of the enterprise groups and their enterprises:

- Group country of global decision centre, defining the country of the enterprise groups
- Group main activity code, defining the complexity of the enterprise groups
- Group employment, defining the size of the enterprise groups
- Number of enterprises with employment in the groups, defining the EU presence of the groups
- Employment of enterprises of the group in the top 5 partner countries

In the most recent EGR experimental statistics publication the following data are published on the groups in the EU and EFTA countries:

- Number of multinational groups operating per country and the number of groups with global decision centre in each country
- Small, medium-sized and large groups by country
- Complexity of enterprise groups by country of group, these data are illustrated with Figure 1 of this document

Figure 1: Complexity of enterprise groups by country of group in 2016 in the most recent EGR dissemination article



Source: EuroGroups Register 2016 data



- European presence of enterprise groups by country of group
- Top 5 EU partner countries in which the groups of country have the highest share of employment
- EU employment by multinational enterprise groups

These data tables are directly extracted from the EuroGroups Register. No other data sources are used. EU-28 aggregates are always published in the tables. Country data are only published in the article where the NSIs agreed to present the data of their groups.

6. Population of groups in the EGR dissemination article

The EGR dissemination article only presents a subpopulation of the yearly EGR frames. For the experimental statistics only the active groups are selected from EGR, those multinational enterprise groups, which have employees in their enterprises in two or more countries, at least one of them in the EU.

In the EGR 2016 frame population 47 621 groups matched this definition; this is 43% of the 111 508 EGR groups in reference year 2016.

57% of the EGR groups did not match the active multinational enterprise group definition and these are not included in the population of the dissemination. The following EGR groups were not considered as active multinational groups and were excluded from the counts given in the experimental statistics article:

- Groups with employment in one country (38% of groups available in EGR), which are composed of:
 - More than one enterprise (28% of all EGR groups)
 - One enterprises only (10% of all EGR groups)
- Groups without employment (19% of all EGR groups), which are composed of:

- More than one enterprise (8% of all EGR groups)
- One enterprise only (9% of all EGR groups)
- No enterprises (2% of all EGR groups) – these group structures are formed by legal units (companies) only.

It is visible, that in large number of EGR group structures there are only legal units (companies); there are no enterprises (statistical representation behind the companies) in one or in multiple countries. The main reason of this can be that the legal unit is empty; the NSIs do not create enterprises for them for the reference year.

In several cases the enterprises in the subsidiary country or countries were created and delivered to EGR, but they have 0 employment or their employment data are missing. Missing employment data of enterprises play a major role not disseminating large number of the EGR groups. The percentage of missing data on enterprise employment provided by the source NSIs varies in EGR from 0 up to 58%. In the future Eurostat will put an emphasis in the EGR process on collecting this variable.

There are further reasons, why some of the groups in the EGR do not match the definition of active multinational enterprise groups. The main remaining one is that some relationships are missing from the database and groups are stored as two or more fragmented groups in the EGR database. The fragments mostly are with few units and with missing data.

7. Way forward

Eurostat aims to switch from production and dissemination of experimental statistics based on EGR data to European official statistics. Therefore, there is a need to investigate user needs and to improve data quality. The first steps already taken have been:

- To work on user feedbacks enriching output tables and taking into account users' interest
- To analyse and improve EGR data quality

Users already expressed their interest in obtaining more detailed statistics. Taking it into account two new figures were disseminated in July 2018 in the first update of the dissemination article. Further discussions with users will be promoted and performed.

Analysis of the population of the active groups and assessment of EGR quality indicators indicate that improvement of EGR micro data is necessary. The main issue observed is on missing relationships and missing employment values. Eurostat in close collaboration with NSIs is working to improve the EGR production and enrich the database, aiming to have a complete picture of the EU part of the groups, covering all their units and characteristics.

Concerning the missing data on employment, Eurostat will monitor and support NSIs to improve data quality. It will be also discussed how data missing in national data sources could be estimated and eventually imputed.

In addition to commercial data Eurostat also considers other data sources to enrich the EGR database, such as web scraping. Eurostat performs projects aiming to retrieve publicly available data from the Internet.

The EGR is a very rich data source, Eurostat and the NSIs are continuously working to improve its quality. Having sufficient quality level in the future, the EGR can be the source of very relevant and valuable European official statistics.