



EUROPEAN COMMISSION
EUROSTAT

Directorate G – Business and trade statistics
G.1 – Coordination and infrastructure development

Luxembourg, 01/12/2023

Methodological note: EuroGroups Register and industry concentration in the European Union

Table of Contents

Methodological note: EuroGroups Register and industry concentration in the European Union ..	1
Background and Context.....	1
Data sources	1
Access to EGR micro data.....	2
Computation.....	3
Breakdowns.....	3
Variables.....	4
Limitations	4

Background and Context

Industry concentration is an interesting topic from the policymaker's point of view. To provide a comprehensive view to the public, Eurostat uses data available from several domains to better understand the industry situation in the European Union. The industry concentration product is a result of the combination of two datasets disseminated by Eurostat. On the one hand, the EuroGroups Register (EGR), which contains the group structure of all multinational enterprise groups that operate within the European Union and the European Free Trade Association territory. On the other, the Structural Business Statistics (SBS) with detailed information about the business economy.

Following up on an exchange of views between the Directorate General Competition of the European Commission and Eurostat, Eurostat's Unit G1-Coordination and Infrastructure Development used the EGR data to compute initial measures of industry concentration. High industry concentration could be an indicator of market failures and trigger some monitoring of the situation.

Data sources

The EGR is the main source of the computation. The EGR is the statistical business register of the EU Member States and EFTA countries for MNE groups. It produces data in yearly cycles and covers microdata on the groups and their enterprises and legal units.

The main value added of the EGR is the possibility to link different enterprises that work under the umbrella of the same group. While in other domains, enterprises are usually treated separately, and thus, as competing players in the industry, with EGR it is possible to treat them together as *working as one* instead of competing. Thus, the aggregated values of the different enterprises of the same group are used. Because enterprises only in one country are out of the scope of the EGR database, a complementary dataset is necessary to properly compute industry concentration indices. Eurostat's Structural Business Statistics describe the business economy through the observation of units engaged in an economic activity; the unit in Structural Business Statistics is the enterprise. An enterprise carries out one or more activities, at one or more locations, and it may comprise one or more legal units. Enterprises that are active in more than one economic activity (plus the value added and turnover they generate, the people they employ, and so on) are classified under the NACE heading corresponding to their principal activity; this is normally the one which generates the largest amount of value added.

SBS contain a comprehensive set of basic variables describing business demographics and employment characteristics, as well as monetary variables (mainly concerning operating income and expenditure, or investment).¹

Access to EGR micro data

The EGR is a statistical business register that serves statistical purposes only. EGR micro data are subject of statistical confidentiality and can only be shared within European Statistical System (ESS) and European System of Central Banks (ESCB). National statistical compilers are given access to all units of the multinational enterprise groups, if at least one of the group's units is within their national territory. These populations can be used for national survey frames, official statistics production, or quality checks.

The necessary data exchange between the national business registers and Eurostat is defined in the [Regulation \(EU\) 2019/2152](#) of the European Parliament and of the Council of 27 November 2019 on European business statistics and in the Implementing [Regulation \(EU\) 2020/1197](#) of 30 July 2020 laying down technical specifications and arrangements pursuant to Regulation (EU) 2019/2152 of the European Parliament and of the Council on European business statistics.

¹ Since 2021, SBS data includes additional NACE Sections than those analysed in the article. However, for the sake of comparability across time (from 2018 to 2020), these have been excluded.

Computation

Industry concentration indices measure the extent to which industry sales are dominated by one or more businesses. At European Union level, one can define an industry using the NACE classification. In this instance, it is measured using the NACE section. The industry concentration can be captured by computing the four-firm concentration ratio, which includes the total sales of the four largest multinational enterprise groups in the industry.

The EGR data presents some limitations in terms of availability of the variable sales; instead, the method used here focuses on the use of the variable of total employment. However, to better understand how appropriate employment as a proxy is, it is important to understand first if there is a strong link between the total sales and the employment figures.

Using the subset of the EGR dataset which contains data for both net turnover (sales) and employment; and split by NACE section, one can compute the Pearson correlation coefficient² at the EU27 to understand how good employment is to measure industry concentration.

The next step is to measure the concentration of the different industries using the available data. For that, it is necessary to obtain the total employment by NACE section of the largest four companies and the total employment by NACE section in the EU27.

The information of the total employment by NACE section can be retrieved from the Structural Business Statistics (SBS), which includes both domestic and multinational enterprises.

Breakdowns

The data is broken down by NACE section available from both domains, the EGR and the SBS. Those are:

- B - Mining and quarrying
- C - Manufacturing
- D - Electricity, gas, steam and air conditioning supply
- E - Water supply; sewerage, waste management and remediation activities
- F - Construction
- G - Wholesale and retail trade; repair of motor vehicles and motorcycles
- H - Transportation and storage
- I - Accommodation and food service activities
- J - Information and communication
- L - Real estate activities
- M - Professional, scientific and technical activities
- N - Administrative and support service activities

² In statistics, the Pearson correlation coefficient measures the linear association between two variables, in this case, employment and net turnover.

Variables

The variables used to compute the industry concentration indices are NACE section of the enterprise, employees (including self-employed) at enterprise level, net turnover at enterprise level, geographical identifiers, and multinational enterprise group identifier of the enterprise:

	EuroGroups Register (EGR)	Structural Business Statistics (SBS)
NACE	ENT_NACE_CODE	NACE_R2
Employment	ENT_PERS_EMPL	V11110 (from 2018 to 2020) / ENT_NR (from 2021 onwards)
Net turnover ³	ENT_TURNNOV	
Geographical area	ENT_COUNTRY_CODE	GEO
Identifiers	GEG_EGR_ID / ENT_EGR_ID	
Other	ENT_STA_CODE	

Limitations

In order to improve the results, and better understand whether or not there is industry concentration at European Union level, it is necessary to improve the coverage of the EGR data, especially for the net turnover variable. Work on that is being carried out by Eurostat and it is considered a priority, even more for the largest multinational enterprise groups, which have a major impact in the economy.

Another limitation of the EGR is that domestic-only groups are excluded from the data. Nevertheless, in practice, one can consider the measures of industry concentration using the EGR data as a lower bound for the actual concentration. This means that, if a domestic-only enterprise group is in one of the largest four companies by employment, then the industry would be even more concentrated than what is computed here.

³ Net turnover refers to the amounts derived from sales after deducting rebates, value added tax, and other taxes linked to turnover.