## How much has national information changed in the Euro Groups Register? Some quality indicators in the scope of the ESBRs project

Esteban Barbado, National Statistical Institute of Spain, ebarbado@ine.es

**Abstract**

*The European System of Interoperable Business Registers (ESS.VIP.ESBRs) is a core Project of the VISION 2020. It aims to create a coordinated infrastructure of frames, serving as backbone for the production of high-quality business statistics in Europe. The Euro Groups Register (EGR) is the heart of the ESBRs system. The EGR contains structural information on multinational groups in the UE. It is managed by a cooperation model based on multiple integrations of microdata provided by Eurostat and the NSIs of European Member States and EFTA countries.*

*The inherent complexity of the EGR business process and the need to reach a good level of stabilization, make the design of active quality policies needed. In fact, a specific Data Quality Program is already in implementation phase. This context will reinforce the role of the EGR as the authoritative frame for globalization statistics in the ESS.*

*This report takes the main elements established in the above mentioned DQP, but using an alternative approach. It provides some quantitative indicators reflecting the changes produced between the Spanish data provided in the EGR production cycle and the information finally registered in the global EGR frame. In addition, the relative importance of the changes produced and their accuracy are also explored, through specific desktop research. The final aim is to provide a wider statistical basis, helping to a better delineation of the priority quality actions for the EGR.*

**Keywords:** European System of Interoperable Business Registers (ESBRs), Euro Groups Register (EGR), Data Quality Program, European Statistical System (ESS)

## The ESBR project and the Euro Groups Register

## The European System of interoperable statistical Business Registers (ESBRs) is a core initiative of the ESS Vision 2020 portfolio. It aims to create a coordinated infrastructure of frames, serving as the backbone for the production of high-quality business statistics in Europe. Because of its importance, an increased number of governance bodies, managers and experts from Eurostat and Member States are involved in its development.

## A continuous improvement of the ESBRs is being undertaken. Several important achievements have already been reported with clear repercussions and benefits; the system of Business Registers is moving towards having more integrated and consistent information on the business units operating in the European Union.

The ESBRs is a network formed by all National Business Registers (NBRs) and the Euro Groups Register (EGR). The EGR is the Statistical Register of the European Communities on multinational enterprise groups, the authentic core of the ESBRs. It includes information on the most influential multinationals operating in the EU and EFTA countries.

The EGR is produced and annually maintained under a strict collaborative model, involving Eurostat and the National Statistical Institutes (NSIs) as the core stakeholders. During the first stages of cycle, the NSIs act as the relevant data providers, according to the information available in the NBRs. Eurostat is in charge of data integration and the generation of a *frozen* system of frames.

Since 2013 some innovations have been progressively incorporated into the production cycle. They were widely implemented for the reference year 2014, which was the starting point of the EGR 2.0 methodology. The definition of more efficient data exchanges, the incorporation of the *authenticity principles[[1]](#footnote-1)* and the new system of identification for the legal units can be pointed out as the main incorporations. All these elements, together with the development of several online applications for data access and interaction, are fostering a better organization of work and more consistent data.

## Quality in Business Registers

Quality is a key requirement for the ESBRs network. Increasing quality of the register’s data means that they more correctly represent the real world to which they refer. Consequently, they will be more intensely used in statistical operations, decision- making and planning.

Since 2015, a common Data Quality Program for NBRs and EGR has been developed, updated and gradually implemented. The adopted approach is based on the *so-called* Deming circle, a tool very commonly used by a wide range of Organizations. Our current program is structured in four components, covering areas directly related to Quality: Reporting, Targets, Assessment and Improvement. This report will focus on Q-Targets component.

The inherent complexity of the EGR model and its strong interconnection with NBRs needs a continuous assessment of the entire production process, related exchange flows and derived outcomes. In this context, Q-Targets represents the main information enabling a complete control of the process. They are normally monitored by a set of indicators, namely:

* *Input* indicators, linked to the formal quality of the national data files
* *Process* indicators, linked to the consolidation of all legal units and the generation of the group´s structures
* *Output* indicators, linked to the consistency of the EGR frames according to the target populations produced by FATS statistics

This report takes the afore-mentioned elements as a reference, but using an alternative approach. It provides some quantitative indicators reflecting the changes produced between the NBR data sent to the EGR and the national data finally registered in the EGR frame. Strictly speaking, this statistical basis does not correspond to input/process/output indicators.

These aggregates add value because they provide elements, which define desirable consistency thresholds between sources. On the other hand, the contribution of third countries to the generation of multinational groups operating in a determined country can also be measured with this new perspective.

This report also examines some discrepancies, following their categorization in terms of relative importance and accuracy. A specific desktop research was carried out by the Spanish BR Unit.

1. **Reference to the sources for analysis**

*3.1 National information*

This concept refers to several files produced from the definitive version of the NBR of reference 2015. The Spanish statistical information on enterprise groups is placed as a module of the national register system, fully consistent with the rest of the statistical units.

The maintenance of the NBR is carried out annually according to a predefined set of input sources. These ones are received, edited, transformed and properly integrated. In the specific case of enterprise groups, a set of commercial, tax and statistical data is used and combined in order to achieve acceptable degrees of coverage.

For this analysis, attention is paid to LEU and REL files (28.869 legal units and 28.814 control relationships) sent during the core part of the EGR cycle. The populations generated and respective selection criteria are:

* National Outward Fats Population OFATS\_NBR. All GDC[[2]](#footnote-2) units of Spanish multinationals plus all their resident subsidiaries*.*
* National Inward Fats Population IFATS\_NBR1. All resident subsidiaries controlled by a non-resident LEU with known identifier of its parent or ultimate unit*.* In the case of a subgroup with non-resident parent and ultimate, both without identifier, the relationships of these subgroups were also sent.
* Complementary National Inward Fats Population IFATS\_NBR2. All resident subsidiaries controlled by a non-resident parent and ultimate, both without identifier*.* Due to this last reason, this population is not sent to the EGR.
* Domestic Groups Population DOM\_NBR. All group heads and the respective subsidiaries of domestic enterprise groups.

OFATS\_NBR and IFATS\_NBR1 conforms the *national information sent to the EGR*, although in several parts of this study, the whole NBR will be used.

*3.2 EGR information*

This source refers to the final frame FINAL\_EGR of reference 2015. It was generated and downloaded on March 2017 through the online EGR\_FATS interface. The national part of this frame corresponds to all multinationals with activities in our country and contains 307,240 legal units operating in 10,197 multinationals.

1. **Comparative analysis and main Indicators**

This study has been conceived in two ways, national information versus EGR and vice versa. At the same time, quantitative data have been produced separately for OFATS (Spanish multinationals with subsidiaries abroad) and IFATS (resident units belonging to foreign multinationals) populations. Due to the size limitations of this paper, we will restrict our results only to one specific population for each sense.

The adopted methodology has been micro-level based. Several record-linkage routines were developed, all of them based on the full coverage and quality of the national IDs for Spanish legal units.

*4.1 National Information versus EGR for IFATS population*

Starting from our NBR, 13,758 resident subsidiaries controlled by a parent or ultimate foreign unit, were sent to the EGR. For all of them, LEID[[3]](#footnote-3) of the controlling unit was provided, as a result of the task “*identification of foreign legal units in the EGR Identification System”*. They represent 96.6% of the whole NBR population. The coverage is even greater in terms of employment and the presence of the most important units is guaranteed.

**Table 1. The national IFATS Population**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **IFATS\_NBR** | **TOTAL** | **IFATS\_NBR1** | **PCT** | **IFATS\_NBR2** | **PCT** |
| **UNITS** | 14,233 | 13,758 | 96.6 | 475 | 3.4 |
| **EMPLOYMENT** | 1,310,712 | 1,300,000 | 99.2 | 10,712 | 0.8 |

Their identification process in the EGR produced the following categories:

* No changes. Units detected with the same profile and the same GDC country
* Changes. Units detected in the EGR but with some changes (IFATS units with different GDC country or belonging to Spanish Multinationals)
* Units not detected in the EGR

The following table illustrates the relative importance of these sub-populations.

 **Table 2. The IFATS population sent; situations in the EGR frame**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **IFATS\_NBR1** | **TOTAL** | **NO CHANGES** | **PCT** | **CHANGES** | **PCT** | **NO EGR** | **PCT** |
| **UNITS** | 13,758 | 10,808 **(a)** | 78.6 | 2,524 | 18.3 | 426 | 3.1 |
| **EMPLOYMENT** | 1,300,00 | 1.044,993 **(a)** | 80.5 | 244,917 | 18.8 | 10,090 | 0.7 |

More than 90% of units with changes continue having the same profile in the EGR but the GDC country has changed. They represent almost 94% in terms of employment. The remaining units are classified in the EGR as subsidiaries to Spanish multinationals. The next table shows how this variable has changed.

 **Table 3. Breakdown of changes for IFATS population**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **IFATS\_NBR1 CHANGES** | **TOTAL** | **NEW PROFILE** | **PCT** | **CHANGE GDC COUNTRY** | **PCT** |
| **UNITS** | 2,524 | 233 | 9.2 | 2,291(b) | 90.8 |
| **EMPLOYMENT** | 244,917 | 15,735 | 6.5 | 229,182(b) | 93.5 |

For the common IFATS population, some *similarity rates* have been calculated, both in terms of units and employment. If we define

$ SR\_{u}=\frac{a}{(a+b)}100$ ; $SR\_{e}=\frac{A}{(A+B)}100$ , we have *SRu* = 82.5 % and *SRe*= 82%.

Units for which the GDC country has changed, are now classified by geographical regions. 61.4% of the Spanish subsidiaries with European control remain in the EGR with the same controlling area.

 **Table 4. Geographical changes of GDC country**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **IFATS\_NBR1 GDC\_COUNTRY** | **TOTAL** | **EU+EFTA** | **JAPAN** | **USA** | **OTHER** |
| **TOTAL** | 2,291 | 1,52 | 26 | 404 | 341 |
| **EU+EFTA** | 1,742 | 1,07 | 23 | 362 | 287 |
| **JAPAN** | 56 | 45 | 0 | 10 | 1 |
| **USA** | 259 | 232 | 2 | 0 | 25 |
| **OTHER** | 234 | 173 | 1 | 32 | 28 |

Several changes experienced in IFATS population have been manually checked, focusing on the biggest units. In the following, a summary of results is presented

*Desktop Research\_1*

* 45 resident legal units were analyzed. Implicitly, these units involved another 48 legal units belonging to the same groups. In summary, the total population investigated comprised 93 I\_FATS units with more than 118,700 employees in Spain
* For this population, 37 foreign Multinationals were identified, their GDC country was analyzed and compared with the registered in both sources
* All cases and decisions made were highly complex. The *correct* GDC country was identified when most of the information found was consistent
* GDC countries registered in the NBR and the EGR normally have some type of close relationship with the multinational
* For some complex cases (Glencore, Airbus, Accenture, .) , it hasn’t been possible to make a well argued and solid decision. It is worth reflecting on the opportunity to define the *nationality* of a global group based on the country concept, which is sometimes not very representative. Some reasons for the non-conclusive results were: several headquarters with multiple functions, a consortium between countries, a tax declaration in one country and the consolidated accounts in another, different nationalities within the boards of directors, etc.
* Quantitatively, we can state that the results seem to be more accurate for EGR**,** although there is no clear predominance
	+ DIRCE provides the correct GDC country for 19 legal units operating in 7 multinationals
	+ EGR provides the correct GDC country for 34 legal units operating in 17 multinationals
	+ For 35 legal units operating in 9 multinationals, the result of the investigation was non-conclusive
	+ Finally, for 5 legal units operating in 4 multinationals, the country code does not coincide with any source

*4.2 EGR versus National Information for OFATS population*

Once EGR data processes were finished, 1,997 Spanish GDCs were included in the final frame. According to the principle of authenticity we can assume that all resident subsidiaries and the respective internal relationships forming their legal structure in the NBR, were reproduced in the EGR.

Consequently, this analysis has been exclusively based on GDCs, evaluating how they are recorded in the NBR. For this purpose, the most complete picture of the national system has been considered. Taking this basis into account, several categories have been identified:

* No changes.- Units detected with the same profile
* Changes.- In this subset, several situations can be found
	+ Belonging to foreign multinationals
	+ Belonging to domestic groups
* Other.- The remaining situations are also considered
* Identified in the NBR as *independent and active* units (not in groups but *alive* units)
* Identified in the NBR as *liquidated* unit
* *Not identified*

The following table summarizes the results obtained

**Table 5. The OFATs population of the EGR; situations in the NBR**

|  |  |  |
| --- | --- | --- |
| **OFATS\_ EGR VERSUS NBR** | **UNITS** | **EMPLOYMENT** |
| **TOTAL** |  | 1.977 | 456.681 |
| **NO CHANGES** | **TOTAL** | 859 **(43.4%)** | 391.642 **(85.5%)** |
| **OFATS\_NBR1** | 859 | 391.642 |
| **OFATS\_NBR2** | 0 | - |
| **CHANGES** | **TOTAL** | 207 **(10.5%)** | 21.983 **(4.8%)** |
| **IFATS\_NBR1** | 10 | 436 |
| **IFATS\_NBR2** | 2 | 515 |
| **DOM\_NB2** | 195 | 21.032 |
| **OTHER** | **TOTAL** | 911 **(46.1%)** | 43.056 **(9.7%)** |
| **INDEPENDENT AND ACTIVE** | 903 | 42.640 |
| **LIQUIDATED** | 8 | 416 |
| **NOT DETECTED** | 0 | - |

The units with different features is highly significant, although less so in terms of size. Nearly 86% of the total employment corresponds to the units with exactly the same profile, which assures an acceptable level of consistency for the most representative units.

Furthermore, 1,547 Spanish GDCs and their respective control perimeters were initially sent to the EGR. These truncated groups had 1,775,138 resident employees. Considering the macro data level, the number of OFATs units in the EGR sharply increases to 27.8%, although this rate is much lower in terms of employees, 6.8%.

Both the differences in aggregates and the discrepancies for common units were investigated and several checks were carried out. A summary of results is provided bellow:

*Desktop Research\_2*

* Considering the global OFATs aggregate, the main reason for the significant variation is due to *actions* of third countries. The incorporation of cross border relationships with resident parent, generates *de facto* Spanish multinationals if this legal unit is not integrated in any other control structure. Generally speaking, two conditions need to be satisfied: Proactive third countries and null or inconsistent national information.
* These situations have been examined in more detail. They correspond to very simple structures, normally formed by two or three legal units with very limited size. A sample was selected for verification and its identification as active OFATs units was practically impossible. For the vast majority of cases we can conclude that the increase of this population really corresponds to false multinationals.
* Regarding the figures presented in the table, the most important units classified in the NBR as *independent and active*, were investigated in terms of employment. Specifically, 86 legal units with more than 50 employees were subject to evaluation:
	+ For 30% of the units, the presence of subsidiaries exclusively out of Spain was verified. This can be explained by a restriction in the design of our national system of groups, which is based on bottom-up relationships. The incorporation of all these structures in the NBR will be considered as a core element for improvement
	+ For 18% of the units, the presence of resident and foreign subsidiaries was confirmed. In a small number of cases, the Spanish subsidiaries were created very recently, after the reference year of the sources
	+ For the remaining units, any evidence of belonging to enterprise groups was found
1. **Final remarks**

This report describes some actions performed by the Spanish NSI in the scope of Quality management linked to the ESBRs system. Their implementation has produced several Q-indicators, aiming to enrich the European initiatives currently adopted.

This statistical basis can also serve as an opportunity to identify new areas for improvement; furthermore the consistency in the network of business registers will be properly evaluated.

The quantification of discrepancies has been complemented with several components of analysis. Among the possibilities, our preferred option focused on identifying the main reasons for their appearance and the study of potential gains in accuracy.

For these purposes, a desktop operation designed inside the BR Unit was carried out. Sources like web pages and other specialized stores of official information were used for data collection. Qualitative information included in financial statements or tax returns was critical for this analysis.

The replication of this experience in other NSIs would be highly desirable. In our view, the presence of heterogeneous results across countries is highly probable, because not all national economies are in the same level of globalization and interactions of third countries can highly influence in the discrepancies

The availability of a set of multi country indicators should be considered as a *good practice*. Its contribution would be useful in many steps of the EGR cycle. Initially, an improved procedure for data extraction in national environments could be designed. In the final steps, this indicators would also serve for a proper delimitation the target EGR Population. Some practical situations to consider:

* New OFATs incorporated in a *destination* country, by third countries sending IFATs data. Defining *false* groups according to prefixed parameters of no relevance
* OFATs sent from an *origin* country and disappeared in the EGR, because countries from the IFATs side do not react. For the most representative cases, Eurostat could acquire the required information on subsidiaries and integrate them in the system.

The discrepancies found in this experience have been, in principle, greater than expected. However, once the reasons of their appearance were identified, a good initiative would be to take joint measures, with the aim of reaching uniform and acceptable consistency rates. Based on this scenario, rates near to 100% could be specifically requested for the TOP European groups or the TOP subsidiaries operating in Europe.

Regarding accuracy, it is more difficult to offer conclusive results. All checks have been closely conditioned by the different situations. Generally speaking, we can state that the collaborative model adds value to the EGR. The data exchange system and its subsequent integration tend to improve the quality of the national information, specifically for the key classification variables. Any deviation from this scenario would justify a reformulation of the current business model.

**References**

ESS Handbook for Quality reports 2014 <http://ec.europa.eu/eurostat/documents/3859598/6651706/KS-GQ-15-003-EN-N.pdf>

ESS.ESBRs project. <http://ec.europa.eu/eurostat/web/ess/esbr>

Directive 2012/17/EU of the European Parliament and of the Council on interconnecting Business Registers. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:156:0001:0009:en:PDF>

COM (2009) 404 - Communication from the Commission to the EP on the production method of EU statistics.<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0404:FIN:EN:PDF>

1. Decision and priority rules adopted in the data consolidation [↑](#footnote-ref-1)
2. Global Decision Centre is the controlling unit of the enterprise group where strategic decisions are made [↑](#footnote-ref-2)
3. The standardized identifier adopted for the legal units in the EGR [↑](#footnote-ref-3)