Enterprise Groups Database - Data manual

Extraction Date - January 2024

BPLIM

2024-10-14

The Enterprise Groups Database is constructed and made available by Banco de Portugal and provides information on the participations in equity capital and voting rights of non-financial corporations operating in Portugal. This dataset contains annual data from 2014 onwards and is mainly based on the information reported through Informação Empresarial Simplificada (IES, Simplified Corporate Information).

Table of contents

General Information

Dataset Designation in English: Enterprise Groups Database

Dataset Designation in Portuguese: Grupos Económicos

Data Type: Longitudinal Data

Unit of Analysis: Participations

Frequency: Annual

Start Date: 2014

Most recent year: 2022

Reference date: The data reports to the end of the fiscal period declared by the firm. For most cases, the fiscal period coincides with the civil year. For those firms with fiscal period different than the civil year, the reference year is the one covering most of the days of the fiscal period.

Data Organization: Data is organized in two files:

- Participation Level File: each row corresponds to a relationship between a participating and an affiliate entity in a given year.
- Ultimate Controlling Entities File: each row corresponds to a firm in a given year.

All files are available in Stata format, version 18.

Version of the Data: The data made available by BPLIM corresponds to a data freeze at a certain time of the year. Therefore, all files contain the information as reported in the extraction date. The most recent update of the data occurred in January 2024.

Languages Available: Variable labels and value labels are available in Portuguese and English. [1]

Related Datasets: BPLIM also makes available the Central Balance Sheet Database and the Central Balance Sheet Harmonized Panel constructed using the information reported through Informação Empresarial Simplificada (IES, Simplified Corporate Information).

Digital Object Identifier: 10.17900/GE.Jan2024.V1

---

[1]To see the labels in English type the following command line in Stata: 'label language en'.

Geographical Coverage

The data refers to firms located in Portugal and the resident and non-resident companies direct or indirectly related to them through participations in equity capital and voting rights.

Population

Enterprise Group Database covers the non-financial corporations operating in Portugal and the resident and non-resident entities related to them through participations in equity capital and voting rights.[2]

Methodology

The Enterprise Groups Database is constructed by the Statistics Department of *Banco de Portugal* mainly using the information on Related Parties and Other Participations reported in the Annex A of *Informação Empresarial Simplificada* (IES, Simplified Corporate Information):

- Table 050603-B: Identification of the entities that **directly** participate in the equity capital of the reporting entity
- Table 050604-B: Identification of the entities in which the reporting entity **directly** participates
- Table 050605-B: Identification of the Ultimate Controlling Entity
- Table 050606-B: Identification of the entities that **indirectly** participate in the equity capital of the reporting entity
- Table 050607-B: Identification of the entities in which the reporting entity **indirectly** participates

This information is reported by resident non-financial corporations. [3] Although all observations in the participation dataset correspond to a direct relationship between two firms (the participating and affiliate entities), the data may also include indirect participations. For example, if the reporting firm A directly participates in the equity capital of firm B that in turn directly participates in the equity capital of firm C, two relationships would be reported: the relationship between firms A and B and the relationship between firms B and C.

The information reported by firms is subject to quality control by the Statistics Department. This control aims to correct reporting inconsistencies and to guarantee that the Enterprise Group is complete. The checks include the elimination of duplicate participations, the identification

---

[2]This dataset does not include natural persons or sole proprietorship. However, for non-resident entities, this may occur to guarantee that the country of the Ultimate Controlling Entity is accurately identified

[3]Tables 050606-B and 050607-B are only reported by resident non-financial corporations identified as the Ultimate Controlling Entity or the controlling entity in Portugal.

of the Ultimate Controlling Entity and the unique identification of non-resident entities. After these checks, every non-resident entity should be assigned with a unique identifier based on the similitude of the Name, Tax Identification Number and Country of origin. The quality control cross checks the information reported in IES with the information published by companies in the Annual Reports and relies on the direct contact with the firm. The quality control also resorts to other complementary sources, such as the Register of Institutions and Affiliates Database (RIAD), the Global Legal Entity Identifier Foundation (GLEIF), the Bank for the Accounts of Companies Harmonized at the firm level (iBACH), and the EuroGroups Register (EGR).

Every year after the quality control process takes place, BPLIM creates a data freeze. [4] The most recent version of the data was extracted on January 2024.

General Definitions

The main concepts used in these data are listed below:

Enterprise Group - Set of firms under the control of the same Ultimate Controlling Entity.

Ultimate Controlling Entity - The company that is at the top of the control chain of the enterprise group. This entity is not direct or indirectly controlled by any other company.

Participating Entity - The entity that owns partly or totally the equity capital or voting rights of other entity(ies).

Affiliate Entity - The entity with all or part of the equity capital or voting rights owned by other entity(ies).

Resident Entity - entity with an economic interest center in the Portuguese economic territory, which may include branches, subsidiaries or a permanent establishment.

Description of Files

The data is organized in two files:

Type of Information	File Name
Participation Level File	GE_A_YFRM_yyyyYYYY_eeee_PART_V01
Ultimate Controlling Entities File	GE_A_YFRM_yyyyYYYY_eeee_UCI_V01

---

[4]After receiving the data, BPLIM excludes the resident entities that are reported due to accounting directives but do not belong to the traditional definition of an Enterprise Group. All participating and affiliate resident entities classified as *Compensation Funds* (*Fundos de Compensação*) are excluded from the data. Additionally, affiliate resident entities classified as *Pension Funds* (*Fundos de Pensões*) or *Mutual Agricultural Credit Banks* (*Caixas de Crédito Agrícola Mútuo*) are also excluded. BPLIM also implements corrections to the data, specifically to the identifiers of foreign companies, to improve the temporal consistency of the information.

Where *A* stands for anonymized and *yyyy* corresponds to the first year available and *YYYY* corresponds to most recent year available. *eeee* reports the extraction date.

Both files contain unique firm identifiers (*tina_participante*, *tina_participada*, *tina*, *tina_uci*) and a reference year (*ano*) . Firm identifiers are anonymized in a consistent manner, which make it possible to merge the information across files and with other firm level datasets made available by BPLIM for resident entities. Labels and value labels were attributed to all categorical variables. All data sets are anonymized.

Description of Variables

Participation Level File

This file is organized by participation and provides detailed information on the relationship between two entities, such as the share of equity capital or voting rights owned by the participating entity. Below we provide a general description of the variables included in the data. For a full account of all variable categories see the "Auxiliary Files" section.

`Reference year` (*ano*) – Reference year of the data.

`Firm identifier` (*tina_participante*, *tina_participada*) - Unique identifier that enables tracking the legal entity over time. This identifier is the anonymized tax identification number (9-digit number) for resident firms and a 11-digit anonymized number for non-resident firms:

   tina_participante - anonymized identifier of the participating entity.

   tina_participada - anonymized identifier of the affiliate entity.

`Country of Origin` (*paisparticipante*, *paisparticipada*):

   paisparticipante - country in which the headquarter of the participating entity is located.

   paisparticipada - country in which the headquarter of the affiliate entity is located.

The country classification is according to the numeric code of ISO 3166. The value labels are available in the "Auxiliary Files".

`Percentage of participation in equity capital` (*per_pd_cs*) – percentage of participation in equity capital.

`Percentage of participation in voting rights` (*per_pd_dirvoto*) – percentage of participation in voting rights.

`Direct investment relationship` (*invdir*) – dummy variable indicating a direct investment relationship between a resident and a non-resident entity. This variable takes value one when

one of the entities designated by direct investor holds at least 10% of the voting rights or equity capital of another entity (direct investment entity).

Ultimate Controlling Entities File

This file provides information on the Ultimate Controlling Entity for firms within the economic groups of Portuguese non-financial corporations. The Ultimate Controlling Entity is mainly identified based on the participations reported by firms. Therefore, an Ultimate Controlling Entity is attributed to all entities with at least one participation of voting rights control (above 50%) in another company. In some cases when it is not possible to identify the Ultimate Controlling Entity, this may be attributed based on the *IES* table identifying this information (050605-B).

Below we provide a general description of the variables included in the data. For a full account of all variable categories see the "Auxiliary Files" section.

`Reference year` (ano) – Reference year of the data.

`Firm identifier` (*tina*) - Unique identifier that enables tracking the legal entity over time. This identifier is the anonymized tax identification number (9-digit number) for resident firms and a 11-digit anonymized number for non-resident firms.

`Country of Origin` (*pais*) - country in which the headquarter of the firm is located. The country classification is according to the numeric code of ISO 3166. The value labels are available in the "Auxiliary Files".

`Ultimate Controlling Entity` (*tina_uci*): anonymized identifier of the entity located at the top of the ownership chain (Ultimate Controlling Entity) of the firm.

`Country of Origin of the Ultimate Controlling Entity` (*paisuci*) - country in which the headquarter of the Ultimate Controlling Entity of the firm is located. The country classification is according to the numeric code of ISO 3166. The value labels are available in the "Auxiliary Files".

Data Limitations

These data sets are constructed based on a mandatory report that may be subject to errors or incompleteness.

Given that the data is constructed based on the report of resident non-financial corporations, it may not be possible to track information regarding participations of foreign and financial entities, which in some cases may not provide a complete picture of the Enterprise Group.

If the participation level information is not accurately reported and a control relationship of voting rights is not identified, then the Ultimate Controlling Entity of the firm may not be identified.

Basic Descriptive Statistics

The number of observations in the files is listed in the tables below.

Table 1- Number of observations in the Participation Level File

```
--------------------------
               |  Frequency
---------------+----------
Reference year |
  2014         |     68,542
  2015         |     71,539
  2016         |     74,833
  2017         |     79,243
  2018         |     84,919
  2019         |     89,661
  2020         |     93,267
  2021         |     99,593
  2022         |    106,735
  Total        |    768,332
--------------------------
```

Table 2- Number of observations in the Ultimate Controlling Entities File

```
--------------------------
               |  Frequency
---------------+----------
Reference year |
  2014         |     42,136
  2015         |     44,134
  2016         |     46,080
```

```
2017            |      48,870
2018            |      51,859
2019            |      54,467
2020            |      56,375
2021            |      59,160
2022            |      62,316
Total           |     465,397
---------------------------
```

Auxiliary Files

For a description of all variables and summary statistics for each data set please check the following auxiliary files: [5]

File	Metafiles	Summary Statistics
Participation Level File	GE_meta_JAN24_PART	GE_metastats_JAN24_PART
Ultimate Controlling Entities File	GE_meta_JAN24_UCI	GE_metastats_JAN24_UCI

Useful Links

- European Business Statistics Manual

Citation of this Dataset

Banco de Portugal Microdata Research Laboratory (BPLIM) (2024). Enterprise Group Database. Extraction: January 2024. Version: V1. Banco de Portugal - BPLIM. Dataset. https://doi.org/10.17900/GE.Jan2024.V1

To cite this data set you can use the package biblatex with the following BibTeX entry:

```
@dataset{GE.Jan2024.V1,
author = {{Banco de Portugal Microdata Research Laboratory (BPLIM)}},
```

---

[5]The summary statistics are available on BPLIM's servers.

```
publisher = {Banco de Portugal – BPLIM. Dataset},
title = {{E}nterprise {G}roup {D}atabase},
year = {2024},
version = {{ V1, Extraction: January 2024}},
doi = {10.17900/GE.Jan2024.V1},
url = {https://doi.org/10.17900/GE.Jan2024.V1}
}
```