

**IFC-ECCBSO-Bank of Spain Workshop on "New insights from financial statements"**

**17 October 2024**

## The Banco de España's Central Balance Sheet Data Office Database: a regional perspective<sup>1</sup>

Alejandro Fernández Cerezo, Borja Fernández-Rosillo  
San Isidro and Natividad Pérez Martín,  
Bank of Spain

---

<sup>1</sup> This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the European Committee of Central Balance Sheet Data Offices (ECCBSO), the Bank of Spain, the BIS, the IFC or the other central banks and institutions represented at the event.

# The Banco de España's Central Balance Sheet Data Office Database: A regional perspective

Alejandro Fernández Cerezo  
Borja Fernández-Rosillo San Isidro  
Natividad Pérez Martín

## Abstract

The availability of a firm-level database that is representative of the productive sector of an economy on an aggregate scale is increasingly important to analyse the heterogeneity of different economic variables at different levels of aggregation (for instance, by region, firm size or sector). This paper seeks, first, to evaluate the representativeness of the Banco de España's Integrated Central Balance Sheet Database (Integrated CBSO database or CBI by its Spanish initials) for conducting regional analysis with firm-level data and, second, to analyse the differences in firm size distribution between the Spanish regions.

Keywords: firm data, firm size distribution, financial reporting.

JEL classification: C81, D21, L11, R11.

## Contents

Abstract.....	1
1. Introduction.....	3
2. The Integrated Central Balance Sheet Database: coverage and representativeness at regional level.....	5
3. Employment: structure and change over time at sectoral and regional level.....	9
4. Regional heterogeneity in firm size distribution.....	13
5. Conclusions.....	21
References.....	22

## 1. Introduction

The availability of a firm-level database that is representative of the productive sector of an economy on an aggregate scale helps analyse the corporate sector's cyclical position and its macroeconomic implications. This is useful when designing public policies that take into account firm heterogeneity in different aspects. A regional analysis using firm-level databases provides an additional layer of detail for a deeper and more accurate understanding of economic developments, particularly in the case of the Spanish economy. First, these sources of firm-level data can be used to explore regional differences across various corporate economic and financial variables, revealing patterns and trends that are not discernible at national level, which can have significant implications for economic policies and regional development strategies. Second, macroeconomic shocks affecting the country as a whole, such as those related to the pandemic (Prades and Tello, 2020) and the recent energy crisis (Fernández Cerezo, Pacce and Sánchez, 2024), may have asymmetric impacts across regions. Also, in countries with a high degree of decentralisation, such as Spain, the regional aspect is especially important, given the existing differences in civil and commercial law and in economic sectors' level of regulation across regions (Mora-Sanguinetti and Soler, 2022).

The Integrated Central Balance Sheet Database (CBI, by its Spanish initials) contains the balance sheets and income statements of a broad range of Spanish non-financial corporations (NFCs) and is therefore a useful source of information for understanding the cyclical position and economic and financial performance of Spain's business sector (Banco de España, 2022a). In particular, it can be used to analyse various aspects, such as firm size, sector of activity and region where the registered office is located.

To provide the general public with easy access to CBI regional data, the Banco de España's Central Balance Sheet Data Office (CBSO) has prepared a report called "**BExplore Regional Central Balance Sheet Data Office**" which provides information on variables regarding the activity, employment, profitability and financial structure of Spanish NFCs, broken down by sector, size and region or province (see Figure 1 for a summary of its content).<sup>1</sup> To this end, information from the annual accounts deposit provided by the Spanish Mercantile Registries and other detailed information provided by CBSO reporting firms was used.<sup>2</sup>

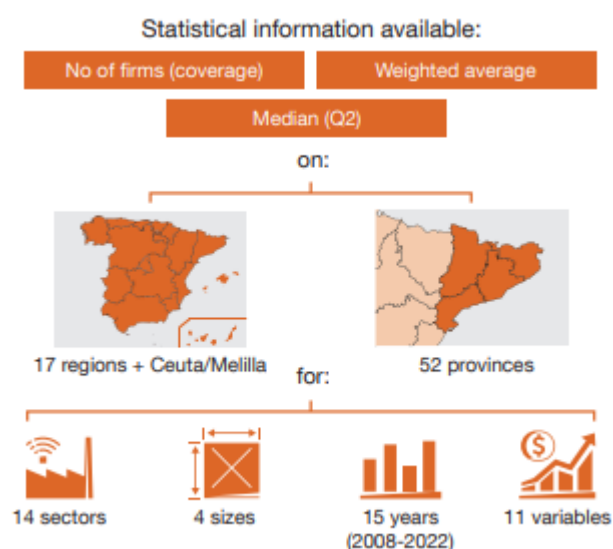
This paper seeks, first, to evaluate the coverage and representativeness, at regional level, of the CBI sample of NFCs and, second, to analyse the cross-regional differences in firm size distribution in Spain. Thus, this paper's main contribution consists of documenting the representativeness of the CBI at regional level. To this end, this paper extends the analysis in Almunia López-Rodríguez and Moral-Benito (2018) to regional level. Among other things, the aforementioned paper showed the

<sup>1</sup> Although the cross-checking of information has been limited to a minimum number of firms and quality control is exhaustive for the purpose of eliminating outliers, the use of the median value (Q2) as an analysis statistic is recommended to compare firms by sector, region and size in the BExplore Regional Central Balance Sheet Data Office tool.

<sup>2</sup> Since 2019 the research community may access CBI firm-level data through the Banco de España's data laboratory (BELab).

CBI's capacity to replicate, at national level, the growth rate of employment in the Spanish private sector and firm size distribution.

*Figure 1. Information available in the report*



*SOURCE: Banco de España*

The results show that the CBI coverage rates (in terms of number of firms) are high in all the regions and that the CBI adequately captures the differences in sectoral structure, employment dynamics and firm size distribution observed in other official sources. As regards firm size, analysis of the CBI's firm-level data shows that there are significant cross-regional differences, even within each sector of activity. This is important, for instance, to understand the factors behind the small size of Spanish enterprises. In this connection, exploiting the CBI's regional dimension allows analysts and academics to address research and economic policy issues of interest for the Spanish economy.

The rest of this paper is structured as follows. Section 2 presents the CBI database and its contents, analysing coverage with respect to the population of firms in each region. Section 3 analyses the capacity of the database to replicate the sectoral distribution of employment across the different regions and changes therein over time. Section 4 analyses cross-regional differences in firm size distribution. Section 5 concludes the paper by exploring some possible future avenues for analysis.

## 2. The Integrated Central Balance Sheet Database: coverage and representativeness at regional level

The CBSO is a service that analyses Spanish non-financial corporations' and corporate groups' economic and financial data. It mainly draws on two sources of yearly information. First, the information obtained directly from firms which voluntarily furnish their data in a specific questionnaire devised by the CBSO (the CBSO Annual Survey or CBA by its Spanish initials).<sup>3</sup> Second, that arising from the cooperation with the Mercantile Registries,<sup>4</sup> which send to the CBSO the annual accounts that firms are obliged to file with them in accordance with financial reporting requirements (the CBB database).<sup>5</sup> Since 1990 firms are required to file their annual accounts with the Mercantile Registry office of the province where their registered office is located.<sup>6</sup> These offices compile and process all the information that is publicly available at the provincial Mercantile Registries and which they ultimately send to the Association of Registrars' Statistical Processes Centre (CPE by its Spanish initials).<sup>7</sup> This is administrative information that is highly reliable, since firms are required by law to provide accurate information about their financial position. In addition, larger firms must undergo an independent review of their annual accounts by a statutory auditor. The Mercantile Registry periodically furnishes the Banco de España with the digitalised raw data (i.e. without filtering out possible accounting inconsistencies or changes in the time series) from the financial statements filed by the firms.

The CBI is constructed through a combination of the two data sources, providing a representative sample of the population of individual NFCs in the Spanish economy, obtained from nearly one million financial statements that are processed yearly. The combined processing of the sources of information on individual firms in the CBI expands the analytical capacity of the database by combining the data available for large firms, which are well represented in the CBA, with those for small and medium-sized enterprises, which are included more comprehensively in the CBB database.

Firms' economic and financial data are processed by the CBSO, harmonising the accounting items over time and cleansing the data provided by the Mercantile

<sup>3</sup> The cooperation with the CBSO consists of annually filling in a questionnaire based on the accounting plan, the content and details of which depend on the firm's size. In 2023 (the latest available year) around 10,000 firms cooperated voluntarily with the CBSO. <https://www.bde.es/wbe/en/areas-actuacion/central-balances/servicios-para-empresas/quien-puede-colaborar/>

<sup>4</sup> See the agreement which defines the cooperation conditions between the Banco de España and the Spanish Mercantile and Property Registrars' Association to support the filing of accounts [https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2023-11942](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2023-11942)

<sup>5</sup> With regard to the filing of accounts, in June 1991 the Banco de España entered into an agreement with the Ministry of Justice for the Madrid Mercantile Registry. The Barcelona, Girona and Tarragona Registries followed suit over the succeeding years. Currently all the provincial Mercantile Registries collaborate with the Banco de España.

<sup>6</sup> There are Mercantile Registries in all the provincial capitals, as well as in Ceuta, Melilla, Ibiza, Mahón, Santiago de Compostela, Puerto de Arrecife, Puerto del Rosario, Santa Cruz de la Palma, San Sebastián de la Gomera and Valverde.

<sup>7</sup> <https://www.registradores.org/registroonline/tramites/mercantil/estadisticas.seam>

Registry to obtain a comparable, consistent and quality database. These data cleansing tasks include, for instance, excluding firms whose accounting year closing date has changed, those providing incomplete information, thus preventing assessment of the units used to prepare the annual accounts (e.g. euro, thousand or million euro)<sup>8</sup> and firms reporting excessive changes in certain items without justifying them. Business combinations between firms, intra-group transactions and the distribution of dividends are appropriately processed to correctly reflect the flows from such operations (for a more detailed analysis, see Banco de España, 2022b).

Table 1 compares, for the 2009-2021 period average,<sup>9</sup> the number of firms comprising the population of NFCs in Spain according to the Central Business Register<sup>10</sup> (column 1), the number of firms whose data are received and processed by the CBSO (column 3) and the number of firms that are suitable for analysis following the data cleansing process mentioned above (column 5). Of the more than 1.25 million firms in the Central Business Register population, the CBSO receives information on almost one million firms, but only around 840,000 are deemed suitable for analysis.<sup>11</sup>

It is noteworthy that the regional distribution of the number of firms fit for analysis in the CBI adequately replicates that of the Central Business Register population, even in the smallest regions. For instance, the number of firms in Rioja, Cantabria and Navarre account for 0.6%, 0.9% and 1.2% of the total population, respectively, (column 2 of Table 1) vs. 0.7%, 0.9% and 1.3% in the CBI (column 6). Among the larger regions, Madrid has a somewhat smaller share in the CBI than in the Central Business Register (18.3% vs. 20.2%), while the opposite is true in Catalonia (19.7% in the CBI vs. 18.8% in the Central Business Register).

<sup>8</sup> For example, firms that fail to report items in the balance sheet or the income statement that are necessary to ensure they are balanced are excluded.

<sup>9</sup> 2008 and 2022 are excluded from the analysis owing to the CBI's lower coverage in those years. Since 2008 was the first year with financial reporting under the new 2007 General Accounting Plan (PGC by its Spanish initials), firms were allowed to provide information relating only to the then current year (2008) and not to 2007. Since the CBSO requires data for two years (current and previous) to cleanse the economic and financial information of each firm, all the firms for which there were no data for 2007 were considered unfit to be included in the CBI database. Also, the information available for 2022 is incomplete, as it is a preliminary sample to be increased in subsequent updates, as a result of the inclusion of new firms. As regards the years prior to 2008, the CBSO has data on 250,000 to 550,000 firms per year from 2000 to 2007. Given the lower coverage for those years, the related historical information will gradually be included in the BExplore Regional Central Balance Sheet Data Office tool using statistical representativeness studies.

<sup>10</sup> Official data on the number of firms in Spain come from the Central Business Register, which is prepared by the National Statistics Institute (INE), which in turn feeds Eurostat's structural business statistics. Included are firms whose legal form falls within the institutional aggregate of NFCs (the CBSO's field of study), mainly public limited companies, private limited companies, general partnerships, limited partnerships and certain civil-law partnerships. Individuals, jointly held property, associations and non-profit private institutions are excluded.

<sup>11</sup> In other words, 16% of firms are rejected on average each year in the period analysed because they do not meet the quality standards set by the CBSO.

*Table 1. Number of firms in the population (Central Business Register) and in the Central Balance Sheet Data Office. Average for the period 2009-2021, by region (a)*

Region	Population of firms, according to the Central Business Register		Firms received in CBB (not adjusted for quality) and CBA		Firms suitable for analysis in the CBI		Coverage rate
	Number of firms	% of national total	Number of firms	% of total firms received	Number of firms	% of total suitable firms	
Region	1	2	3	4	5	6	7 (5 / 1)
Andalusia	176,041	14.0	133,981	13.4	110,964	13.3	63.0
Aragon	33,341	2.7	29,733	3.0	25,544	3.1	76.6
Asturias	19,964	1.6	17,303	1.7	14,967	1.8	75.0
Balearic Islands	36,188	2.9	32,020	3.2	26,478	3.2	73.2
Canary Islands	52,143	4.1	32,654	3.3	27,540	3.3	52.8
Cantabria	11,280	0.9	8,628	0.9	7,421	0.9	65.8
Castile-Leon	53,297	4.2	43,978	4.4	37,489	4.5	70.3
Castile-La Mancha	46,706	3.7	38,398	3.8	32,799	3.9	70.2
Catalonia	235,921	18.8	196,928	19.7	164,825	19.7	69.9
Valencia	139,601	11.1	114,250	11.4	97,495	11.6	69.8
Extremadura	18,308	1.5	14,649	1.5	12,336	1.5	67.4
Galicia	68,473	5.4	55,226	5.5	47,801	5.7	69.8
Madrid	254,415	20.2	187,661	18.8	152,894	18.3	60.1
Murcia	35,859	2.9	28,410	2.8	23,946	2.9	66.8
Navarre	15,615	1.2	12,724	1.3	10,681	1.3	68.4
Basque Country	50,946	4.0	44,648	4.5	38,141	4.6	74.9
Rioja	7,571	0.6	6,679	0.7	5,726	0.7	75.6
Total	1,257,924	100.0	997,870	100.0	837,046	100.0	66.5

*SOURCES: Banco de España and INE*

*(a) Including firms whose legal form lies within the scope of analysis of the Central Balance Sheet Data Office. Excluding individuals, jointly held property and associations. Including firms with zero or more employees.*

Another important aspect for demonstrating the CBI's contents is to analyse the coverage rate, defined as the ratio of the number of NFCs fit for analysis in the CBI to the number included in the Central Business Register (column 7). The coverage rate at national level is 66.5% on average for the period 2009-2021, exceeding 70% in Aragon, the Balearic Islands, Rioja, the Basque Country and Asturias, and is somewhat lower in the Canary Islands and Madrid (52.8% and 60.1%, respectively). These cross-regional differences in coverage are mainly due to provincial differences in the reception, processing and digitalisation of the annual accounts filed with the Mercantile Registries. As a result, the information available in the Association of Registrars' CPE (which the CBSO draws from) changes by region. In addition to the percentage of firms that continue to file in print (approximately 8% of the total for Spain, with some regional differences), some annual accounts are filed late, whether in breach of the law or because the firm was unable to approve its accounts at the Annual General Meeting (a shortcoming that is usually remedied during a subsequent filing).<sup>12</sup>

Table 2 analyses the changes over time in the CBI's coverage rates. An upward trend is observed over the period analysed across most regions. This is the result of, first, an ongoing increase in the number of firms filing their annual accounts with the Mercantile Registries, as well as the greater use of the XBRL electronic format (reducing the filing in print) and, second, an improvement in the CPE's

<sup>12</sup> Other factors causing regional coverage differences are: enterprises accounted for as population but which do not file their accounts with the Registries, such as civil-law partnerships, firms that are registered but which are dormant and firms that do not meet the information quality criteria and, consequently, are not included in the CBSO publications.



information receipt processes, granting the CBSO greater access to the Mercantile Registry databases, thanks to the cooperation agreements entered into by the two institutions.

*Table 2. Coverage rate (number of NFCs included in the CBI relative to the Central Business Register), by region (a)*

Region	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Average (2009-2021)
Andalusia	51.6	52.1	54.9	58.3	60.9	64.1	65.1	64.3	61.8	59.2	65.6	62.2	61.9	63.0
Aragon	67.9	72.4	73.4	73.3	73.6	73.5	74.9	74.9	71.3	72.4	78.1	78.1	78.8	76.6
Asturias	69.3	67.7	70.8	71.8	71.8	73.7	74.3	73.0	71.5	70.6	76.3	74.1	75.0	75.0
Balearic Islands	65.0	67.7	73.1	72.5	70.9	75.0	75.1	74.5	61.5	66.7	73.8	72.9	70.7	73.2
Canary Islands	50.9	50.0	51.4	50.5	51.9	53.2	55.1	52.6	48.1	40.8	47.4	56.0	44.4	52.8
Cantabria	57.1	52.6	52.7	55.4	59.6	66.2	73.5	71.5	63.7	63.3	69.1	68.4	73.1	65.8
Castile-Leon	64.7	64.7	65.1	67.2	67.4	68.9	70.4	70.8	63.9	66.0	70.1	70.8	71.6	70.3
Castile-La Mancha	61.6	61.6	64.6	65.9	67.1	69.6	71.6	70.6	66.9	66.9	72.4	73.0	66.8	70.2
Catalonia	61.3	61.6	63.4	65.8	70.2	71.1	69.7	68.8	64.6	63.5	67.6	67.6	68.4	69.9
Valencia	62.2	61.8	63.6	65.1	66.5	69.3	69.3	70.1	67.6	66.8	71.6	70.7	72.5	69.8
Extremadura	62.8	62.8	64.7	60.6	60.9	67.3	68.4	69.3	60.1	60.5	62.5	66.7	69.4	67.4
Galicia	63.3	63.1	66.7	66.7	67.1	67.1	68.5	70.2	63.7	64.9	72.1	72.8	73.7	69.8
Madrid	50.7	51.8	53.5	55.0	56.5	58.6	58.3	58.1	58.4	60.6	63.1	62.0	61.5	60.1
Murcia	57.3	58.1	59.9	61.8	62.9	63.8	65.6	66.3	61.4	61.8	68.9	69.9	70.6	66.8
Navarre	57.6	57.7	59.6	61.2	66.3	67.5	69.6	67.7	66.0	67.9	72.3	69.3	67.8	68.4
Basque Country	58.3	56.3	62.3	67.0	72.1	76.2	73.9	76.8	76.1	77.0	79.0	76.7	75.8	74.9
Rioja	65.8	64.7	69.0	71.4	72.3	73.5	73.2	74.7	72.9	77.0	80.0	79.6	77.7	75.6
Total	57.9	58.2	60.5	62.2	64.2	66.3	66.5	66.3	63.1	63.0	67.6	67.2	66.8	66.5

#### *SOURCES: Banco de España and INE*

*(a) Including firms whose legal form lies within the scope of analysis of the Central Balance Sheet Data Office. Excluding individuals, jointly held property and associations. Including firms with zero or more employees.*

A very important aspect when analysing the CBI's firm-level data from a regional viewpoint is that each firm is located in the province where it is headquartered, regardless of whether it has production sites in other areas. This is because there is only one balance sheet, which cannot therefore be fractioned on the basis of the percentage of sales or employment by province or region. This aspect is vitally important for any comparison of the CBI results at regional level with other sources of aggregate information.

It is worth asking whether allocating the firms' annual accounts in full to the region where they are headquartered skews the CBI's regional analysis. In order to approximate the geographical location of their activities, the survey for CBA reporting firms includes a question about the regional distribution of personnel costs.<sup>13</sup> In particular, a sub-sample of around 2,800 large firms report the percentage of personnel costs incurred in each region. Based on this information it can be deduced that firms being located in several regions could have a limited impact on

<sup>13</sup> This information is not available for CBSO users for confidentiality reasons.

the CBI's regional representativeness, since 87% of the firms in this sub-sample incur over 90% of their personnel costs in the same region where their headquarters are located. This means that the activity of the majority of firms that provide detailed information of their labour costs is concentrated in a single region. Also, only 5% of them have more than 50% of their personnel costs in regions other than where they are headquartered.

In sum, the CBI sample available is reasonably representative of the population of NFCs at regional level.

### 3. Employment: structure and change over time at sectoral and regional level

This section compares the sectoral distribution of the CBI by region with that from official sources, in order to assess the CBI's ability to capture cross-regional differences in productive systems. We are thus able to examine whether the composition of the CBI is skewed towards a certain set of firms belonging to one sector. This is of particular interest, for example, for any research aiming to take into account the asymmetric effects of economic shocks on sectors and regions.

The analysis focuses on the non-financial and non-agricultural market economy, i.e. we disregard firms operating in the following sectors: i) public administration and defence; compulsory social security, education, health and social work; ii) financial and insurance activities; iii) agriculture, forestry and fishing; and iv) activities of households as employers and producers and activities of membership organisations.<sup>14</sup>

To assess the representativeness of the CBI when capturing regional differences in sectoral specialisation, we use as reference the sectoral distribution of the employees registered under the General Social Security Regime.<sup>15</sup> It should be noted that, to calculate the number of employees in each sector and region, Social Security uses a slightly different method from that used in financial statements. The former provides the average number of workers each month, unadjusted for temporary and part-time employees. By contrast, the latter report the average

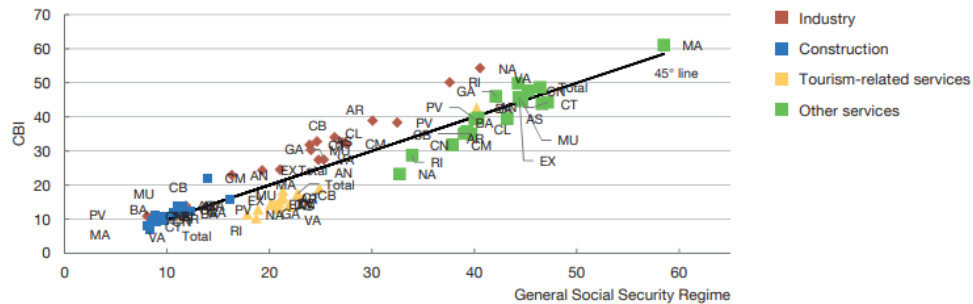
<sup>14</sup> Observations with no employees are removed because these firms mostly represent dormant firms or firms incorporated for tax purposes. The self-employed do not fall within the CBI's scope.

<sup>15</sup> The General Social Security Regime includes firms' employees under that regime. Therefore, those registered under the Special Social Security Regime for the Self-Employed are excluded, as the CBI does not include natural persons. [https://w6.seg-social.es/PXWeb/pxweb/es/Afiliados%20en%20alta%20laboral/Afiliados%20en%20alta%20laboral\\_\\_Afiliados%20Medios/04mb.%20Por%20CC.AA,%20%20seccion%20de%20actividad%20y%20regimen%20\(RG%20y%20RETA\).px/](https://w6.seg-social.es/PXWeb/pxweb/es/Afiliados%20en%20alta%20laboral/Afiliados%20en%20alta%20laboral__Afiliados%20Medios/04mb.%20Por%20CC.AA,%20%20seccion%20de%20actividad%20y%20regimen%20(RG%20y%20RETA).px/)

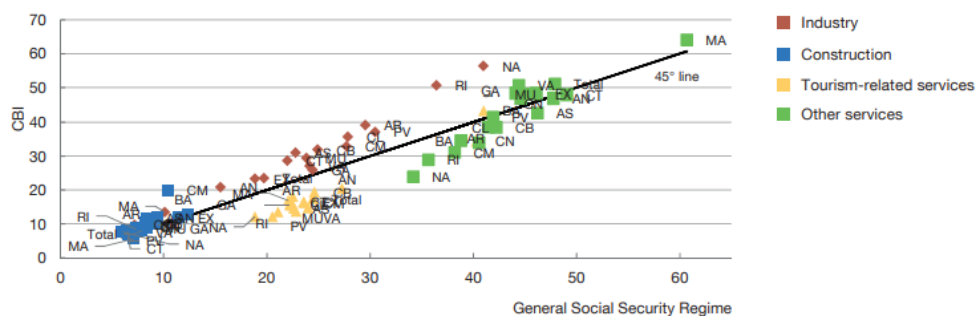
number of full-time equivalents throughout the year, considering temporary employees as a fraction of full-time employees.<sup>16</sup>

*Chart 1. Sectoral distribution of employment, by region (% of total employment in each region) (a)*

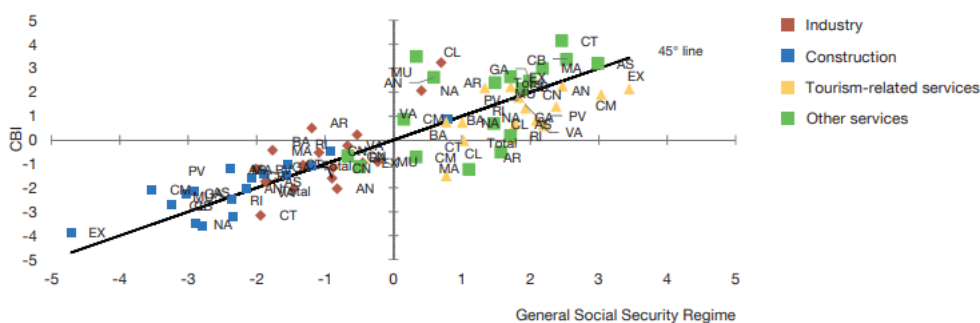
*1.a Average for the period 2009-2014 (%)*



*1.b Average for the period 2015-2021 (%)*



*1.c Change between the averages for the periods 2009-2014 and 2015-2021*



**SOURCES:** *Banco de España*

(a) "Industry" includes the firms under NACE Rev. 2 sections B (mining and quarrying), C (manufacturing), D (electricity, gas, steam and air-conditioning supply) and E (water supply, sewerage, waste management and remediation). "Construction" includes firms in sector F (construction). "Tourism-related services" includes firms

<sup>16</sup> The CBI employment figure reflects the hires and terminations of employees reported throughout the year, weighted by the months they worked at the firm. To convert part-time, temporary and permanent seasonal employment into full-time equivalent employment, the hours worked by all employees with these types of contracts are added together and then divided by 1,826 hours (the assumed number of hours that a full-time employee works per year).

*in sectors H (transportation), I (accommodation and food service activities) and R (arts, entertainment and recreation). "Other services" includes firms in sectors G (wholesale and retail trade), J (information and communication), M (professional, scientific and technical activities) and N (administrative and support service activities). AN: Andalusia, AR: Aragon, AS: Asturias, BA: Balearic Islands, CN: Canary Islands, CB: Cantabria, CL: Castile-Leon, CM: Castile-La Mancha, CT: Catalonia, VA: Valencia, EX: Extremadura, GA: Galicia, MA: Madrid, MU: Murcia, NA: Navarre, PV: Basque Country, RI: Rioja.*

Since the sample period (2009-2021) includes a recession and a recovery, it has been divided in two: Chart 1.a depicts the average share of employment according to the CBI and Social Security in four major sectors<sup>17</sup> and by region for the period 2009-2014, while Chart 1.b does the same for the period 2015-2021. The results of this comparison show that the CBI adequately reflects the differences in the regional distribution of employment observed in the social security registrations data. With regard to the services sector, the Balearic Islands and the Canary Islands have the highest share of employment in tourism-related services according to the CBI (around 42% and 32%, respectively). This is mirrored in the social security registrations data (around 40% and 37%, respectively). Meanwhile, Madrid is the region with the highest share of employment in other services (61% in the CBI and 59% in the social security data). Turning to the secondary sector, Navarre and Rioja are the two regions with the highest share of employment in industry in the two databases. However, industrial employment is overrepresented in the CBI in all regions. Castile-La Mancha and Extremadura are the regions with the highest share of construction employment in both the CBI and the social security registrations data.

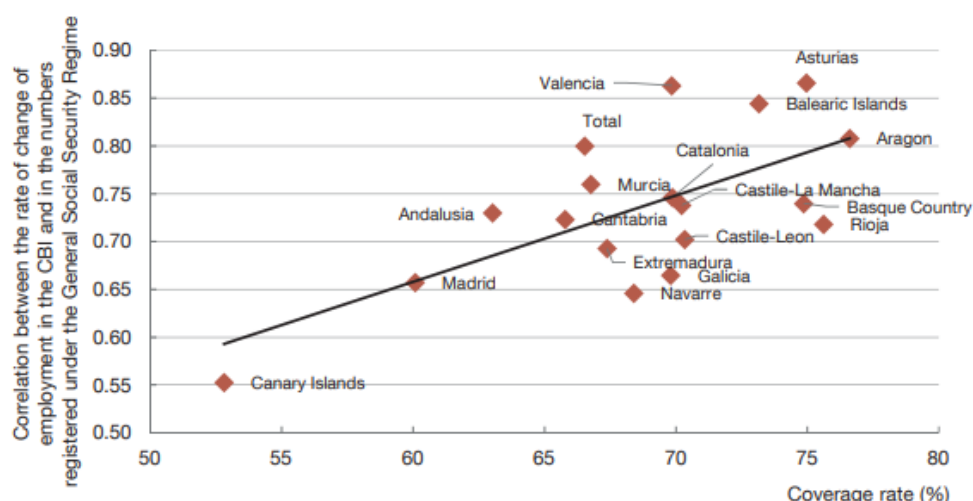
The CBI also adequately captures the change in the sectoral composition of employment that took place in the two periods under review. At national level, the shift to a services economy stepped up, with the percentage of industrial and construction employment falling by 1 pp and 1.4 pp, respectively, and that of services employment increasing by 2.4 pp, according to the CBI sample (see Chart 1.c). These figures are similar to those obtained from the social security data, which show a fall in employment in industry and construction of 1.2 pp and 1.8 pp, respectively, and a 3 pp rise in the share of services employment. By region, the CBI also performs notably when reflecting this sectoral shift. For example, according to the social security registrations data, the share of industrial employment decreased in all regions apart from Navarre, Castile-Leon and Castile-La Mancha. This is also reflected in the CBI data. Extremadura saw the largest decline in the share of construction employment in both the CBI and the social security registrations data (3.9 pp and 4.7 pp, respectively). The increase in the share of tourism-related services employment in the Canary Islands (2.3 pp) is notable. The official data yield a similar figure (2.5 pp). Catalonia, Asturias and Cantabria are the three regions where the share of

<sup>17</sup> "Industry" includes the firms under NACE Rev. 2 sections B (mining and quarrying), C (manufacturing), D (electricity, gas, steam and air-conditioning supply) and E (water supply, sewerage, waste management and remediation). "Construction" includes firms in section F (construction). "Tourism-related services" includes those firms in sections H (transportation and storage), I (accommodation and food service activities) and R (arts, entertainment and recreation). "Other services" includes firms in sections G (wholesale and retail trade), J (information and communication), M (professional, scientific and technical activities) and N (administrative and support service activities).

employment in other services has gained the most both in the CBI and in the social security data.

Lastly, we assessed the CBI's ability to replicate the dynamics of aggregate employment at regional level throughout the period 2009-2021. By focusing on dynamics rather than levels, this exercise is particularly relevant for any research with a macroeconomic perspective. The CBI is highly accurate at reproducing the dynamics of overall employment in Spain in the period 2000-2013 (Almunia, López-Rodríguez and Moral-Benito, 2018). The CBI also proves to be highly accurate when performing this analysis for the reference period of this paper (2009-2021): at national level, the correlation between the annual rate of change in total employment in the CBI and that of the General Social Security Regime data is 0.82.<sup>18</sup> When performing the same exercise for total employment in each region, the correlation between the CBI and the social security time series is 0.72 on average for the 17 regions, and tends to be higher in those regions where the CBI has a higher coverage rate (such as Asturias, Aragon and the Balearic Islands) and lower where its coverage rate is lower (e.g. the Canary Islands) (see Chart 2).

*Chart 2. CBI coverage rates and correlation of employment dynamics between the CBI and registration under the General Social Security Regime (a)*



*SOURCES: Banco de España, INE and Ministerio de Inclusión, Seguridad Social y Migraciones*

*(a) Drawing on annual employment data for the period 2009-2021. Including non-financial and non-agricultural market economy firms, excluding firms without any employees. The coverage rate is defined as the ratio between the number of NFCs suitable for analysis in the CBI and the number of firms in the Central Business Register*

<sup>18</sup> In Almunia, López-Rodríguez and Moral-Benito (2018), the correlation between total employment in the CBI firm-level sample and employment in the National Accounts is 0.91 for the period 2000-2013.

In sum, broadly speaking, the available CBI sample adequately replicates the sectoral distribution of employment in the different regions and the changes therein over time in aggregate terms.

## 4. Regional heterogeneity in firm size distribution

One of the Spanish business sector's main characteristics is the predominance of smaller enterprises, which is key to understanding Spain's low aggregate productivity (Moral-Benito, 2018).<sup>19</sup> It is also at smaller enterprises where a wider negative productivity gap vis-à-vis other European countries is observed.<sup>20</sup>

This section analyses the existing cross-regional differences in firm size distribution. First, we show that the CBI adequately replicates the distribution of firm size and employment at national level. Second, we document the distribution of the number of firms and employment in the CBI by size group and region, with particular focus on the enterprises with fewer than 50 employees. Lastly, we illustrate the changes over recent years in the distribution of employment in the CBI, by firm size and region.

Prior to analysing the regional differences in firm size, we should examine whether the CBI sample adequately reproduces the distribution of the number of firms and employment by firm size group reported in the official statistics at national level. Specifically, in terms of the number of firms, we use as reference the distribution of active firms by number of employees and legal form<sup>21</sup> reported by the Central Business Register,<sup>22</sup> while for employment we use the Statistics for Social Security-registered Firms (*Estadística de Empresas Inscritas en la Seguridad Social*). As in the previous section, the analysis focuses on the non-financial and non-agricultural market economy, and excludes firms with zero employees.

Chart 3.a shows that the distribution of the CBI sample of firms is considerably similar to that of the population in the period 2009-2021. Specifically, large firms (250 or more employees) and medium-sized enterprises (50-249 employees) accounted for 2% and 0.5%, respectively, of the total sample in both databases. However, microenterprises (1-9 employees) are slightly underrepresented in the CBI (83.3% of firms in this database, versus 84.9% in the Central Business Register), while small enterprises (10-49 employees) are slightly overrepresented (14.2% of firms in the CBI, versus 12.6% in the Central Business Register). In terms of the distribution of employment, Chart 3.b shows that the average percentage of workers at large firms in the period 2013-2021 is higher in

<sup>19</sup> When discussing firm size and its implications, the evidence suggests that the causal direction is productivity to size, not size to productivity. In other words, the most efficient and productive firms tend to gain market share and grow, while the least productive tend to lose market share.

<sup>20</sup> Banco de España (2024) and Gavilán and García Coto (2024).

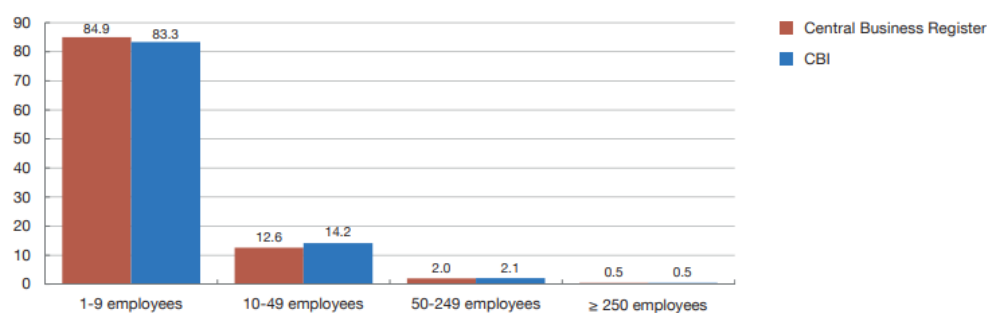
<sup>21</sup> We analyse firms incorporated as legal persons in the form of public limited companies, private limited companies and cooperatives.

<sup>22</sup> To define the size groups, the Central Business Register considers the total number of employees at year-end, without adjusting for temporary and part-time employees.

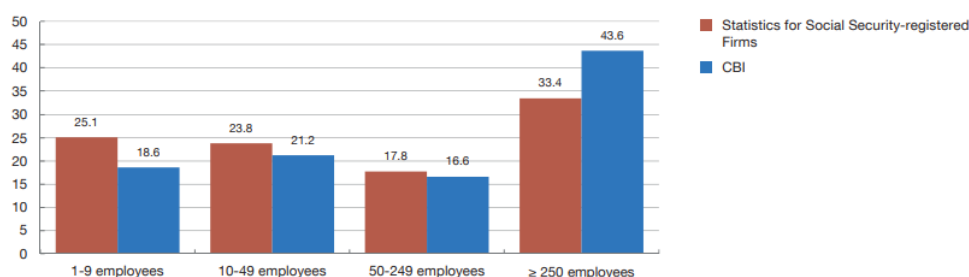
the CBI (43.6%) than in the social security data (33.4%).<sup>23</sup> By contrast, in the CBI the percentage of employment in the other size categories is lower than in the social security data, especially in the microenterprise segment (18.6% in the CBI, versus 25.1% in the social security data). These findings are in line with Almunia, López-Rodríguez and Moral-Benito (2018), who show that the left tail of the distribution of firms has a lower coverage rate in the CBI, due to poorer data quality in this size category, which stems, for example, from a higher relative concentration of missing data in the employment variable.

*Chart 3. Distribution of the number of firms at national level, by firm size (a)*

*3.a Percentage of firms, by firm size. Average for the period 2009-2021*



*3.b Percentage of employment, by firm size. Average for the period 2013-2021*



*SOURCES: Banco de España, INE and Ministerio de Inclusión, Seguridad Social y Migraciones*

*(a) Including non-financial and non-agricultural market economy firms, excluding firms without any employees*

In light of the CBI's representativeness when reflecting the national firm size distribution, next we examine the same distribution in the CBI, but at regional level. In terms of the number of firms, a CBI-Central Business Register comparison of the regional distribution of this variable by firm size is not possible, because the Central Business Register does not provide a breakdown of the population of firms by number of employees, legal form and region. In any event, Table 3.a highlights some cross-regional differences in firm size distribution. First, the Basque Country and Rioja have the lowest percentage of microenterprises (under 80%), while Castile-La Mancha, Extremadura and Castile-Leon have the highest proportion of these firms (over 85%). Second, Madrid stands out as the region with the highest percentage of

<sup>23</sup> The first year available with information broken down by sector and size in the Statistics for Social Security-registered Firms is 2013.

large firms (1.1%), which could be related to the “headquarters effect”, which consists of firms – especially the largest ones – tending to locate their headquarters in major cities in order to harness the wealth of highly specialised professional profiles and the concentration of certain services.<sup>24</sup>

*Table 3. Distribution of the number of firms and employment in the CBI, by firm size (a)*

### *3.a Distribution of the number of firms, by firm size*

Region	1-9 employees	10-49 employees	50-249 employees	≥ 250 employees
Andalusia	84.8	13.4	1.6	0.3
Aragon	82.9	14.8	2.0	0.4
Asturias	83.5	14.3	1.9	0.4
Balearic Islands	83.3	14.8	1.6	0.3
Canary Islands	80.3	16.6	2.7	0.4
Cantabria	80.5	16.8	2.2	0.5
Castile-Leon	85.3	13.1	1.4	0.2
Castile-La Mancha	86.0	12.5	1.3	0.2
Catalonia	82.5	14.6	2.4	0.5
Valencia	84.3	13.6	1.8	0.3
Extremadura	85.5	13.0	1.4	0.2
Galicia	85.2	12.8	1.7	0.3
Madrid	82.6	13.7	2.7	1.1
Murcia	81.9	15.6	2.2	0.4
Navarre	80.5	15.9	3.0	0.6
Basque Country	79.4	17.3	2.7	0.6
Rioja	79.7	17.9	2.2	0.2
Total	83.3	14.2	2.1	0.5

### *3.b Distribution of employment, by firm size*

Region	1-9 employees	10-49 employees	50-249 employees	≥ 250 employees
Andalusia	30.3	31.3	18.3	20.2
Aragon	24.9	27.9	18.6	28.6
Asturias	23.7	25.3	17.3	33.8
Balearic Islands	25.5	29.3	15.4	29.7
Canary Islands	21.9	29.3	24.3	24.5
Cantabria	23.6	30.9	20.7	24.7
Castile-Leon	31.3	30.1	16.2	22.4
Castile-La Mancha	34.5	31.9	17.2	16.3
Catalonia	19.7	23.9	19.5	36.9
Valencia	24.5	26.1	17.6	31.8
Extremadura	36.8	32.8	17.2	13.1
Galicia	25.1	24.5	16.2	34.3
Madrid	8.7	10.5	10.8	69.9
Murcia	23.0	28.6	20.5	27.9
Navarre	19.6	26.2	23.4	30.9
Basque Country	17.8	25.0	19.6	37.6
Rioja	27.6	38.7	22.6	11.2
Total	19.0	21.6	16.1	43.3

**SOURCE:** *Banco de España*

(a) *Average for the period 2009–2021. Including non-financial and non-agricultural market economy firms, excluding firms without any employees*

<sup>24</sup> However, when considering firms overall, the coverage rate of the CBI in Madrid is lower than the national average (see Tables 1 and 2), suggesting that this region’s lower representation is concentrated in the small enterprise category.



Turning to employment, Table 3.b shows the distribution of the number of employees by firm size according to the CBI. As was the case when analysing the number of firms, the regional distribution of employment by firm size cannot be compared with the Statistics for Social Security-registered Firms, as they do not provide information broken down by size or region. At national level, while firms with fewer than 50 employees accounted for 97.4% of the total on average in the period 2009-2021, they only made up 40.6% of employment.<sup>25</sup> Meanwhile, large firms accounted for just 0.5% of the total, but for 43.3% of employment. However, there are considerable cross-regional differences. First, the share of employment at large firms (250 or more employees) is almost 70% in Madrid, virtually doubling the proportion in the Basque Country (37.6%) and Catalonia (36.9%). Second, at the other end of the scale, in Castile-La Mancha and Extremadura microenterprises (1-9 employees) account for 34.5% and 36.8% of employment, respectively. Turning to the intermediate size categories, Rioja is the region with the most employment at small enterprises (38.7%), while the Canary Islands and Navarre are the regions with a higher share of employment at medium-sized enterprises (24.3% and 23.4%, respectively).

Against this backdrop, the CBI is a unique source in Spain for analysing regional heterogeneity in firm size distribution since other sources, such as the Central Business Register and the Statistics for Social Security-registered Firms, do not provide a breakdown of employment by firm size, sector and region. Accordingly, the CBI allows for analysis of the Spanish business sector from new dimensions, some of which are considerably important from an economic policy standpoint. By way of example, the remainder of this section uses the CBI to study the regional patterns of the concentration of employment at smaller enterprises. This is a matter that, as mentioned above, at national level constitutes one of the most important factors behind the low productivity growth in the Spanish economy.

In order to isolate the sectoral composition effect, Table 4 shows the share of employment concentrated at smaller enterprises (1-49 employees), by region and across all four sectors analysed.<sup>26</sup> At national level, the concentration of employees at smaller enterprises is higher in construction (66.8%) and in tourism-related services (47.2%), while in industry (36.1%) and in other services (36%) there is a lower share of employment in this size category. This indicates that regional differences in productive specialisation affect firm size distribution. In addition, within each sector there are also notable cross-regional differences in firm size, which, as mentioned above, can only be analysed thanks to the richness of the CBI firm-level data. For example, in industry, Asturias, Navarre and Madrid are the regions with the lowest share of employment at smaller enterprises (30.7%, 26.5% and 19%, respectively),

<sup>25</sup> These aggregate national figures are similar to those drawn from the Statistics for Social Security-registered Firms, according to which smaller enterprises accounted for 48% of employees in the period 2013-2021.

<sup>26</sup> The 50 employee threshold is important because academic literature has shown that the existence of rules and regulations that are applied to firms with 50 or more employees (such as the requirement to establish a workers' committee) may discourage business growth and lead to an over-allocation of resources to less productive firms. Maza (2024) examines how thresholds in accounting and financial reporting requirements influence businesses' growth decisions. Almunia, Jimeno, López-Rodríguez and Petit (2024) analyse the effects of the interaction of labour and fiscal regulations on the size of Spanish firms.

while in the Balearic Islands it is almost 70% and in Extremadura it stands at close to 60%.

*Table 4. Share of employment at smaller firms (1-49 employees) in the CBI (%) (a)*

Region	Industry	Construction	Tourism-related services	Other services	Total
Andalusia	50.0	74.5	70.1	59.8	61.6
Aragon	35.0	82.9	59.8	61.3	52.9
Asturias	30.7	56.3	77.3	51.0	49.0
Balearic Islands	67.4	84.3	40.6	58.7	54.9
Canary Islands	52.0	70.7	44.2	52.2	51.2
Cantabria	37.2	74.4	66.7	57.5	54.5
Castile-Leon	42.1	87.8	83.0	62.2	61.4
Castile-La Mancha	55.5	58.4	75.3	78.8	66.5
Catalonia	33.2	73.6	50.5	43.1	43.6
Valencia	49.4	82.0	64.7	42.2	50.7
Extremadura	58.5	87.7	82.5	65.8	69.6
Galicia	39.3	76.1	63.8	44.7	49.6
Madrid	19.0	40.2	22.9	16.3	19.3
Murcia	43.0	78.1	64.8	47.7	51.6
Navarre	26.5	72.7	67.1	69.6	45.7
Basque Country	33.0	64.5	54.3	43.7	42.9
Rioja	50.7	94.0	80.9	78.6	66.3
Total	36.1	66.8	47.2	36.0	40.6

*SOURCE: Banco de España*

*(a) Average for the period 2009-2021. Including non-financial and non-agricultural market economy firms, excluding firms without any employees.*

Turning to construction, according to the CBI around 90% of workers in Rioja, Castile-Leon and Extremadura are employed at smaller enterprises, versus the lowest percentages in Asturias (56.3%), Castile-La Mancha (58.4%) and Madrid (40.2%). Meanwhile, in tourism-related services, the highest percentage of employment at smaller enterprises exceeds 80% in Extremadura and Castile-Leon, in contrast to the Canary Islands and the Balearic Islands (around 40%). With regard to other services, the high share of employment at smaller enterprises in Rioja and Castile-La Mancha (almost 80%) contrasts with that in Madrid (16.3%) and Catalonia and Valencia (around 40%).

Given this high regional heterogeneity in the distribution of employment by firm size, we should consider the factors that are potentially behind these differences. In this regard, the first aspect to consider is that, broadly speaking, the share of employment at smaller enterprises tends to be lower in those sectors in which the region is more specialised (measured by the sector's share of total employment in each region), except for construction, where this correlation is not observed (see Chart 4).

The CBI firm-level data also allow us to examine whether the regional firm size distribution has changed over time. Indeed, one of the most notable developments in the Spanish economy in recent years has been the trend towards a lower concentration of employment at smaller enterprises. Specifically, according to the Statistics for Social Security-registered Firms, the percentage of employees at such firms fell from 43.3% in 2009 to 42.2% in 2021 and 39.5% in 2024.

The CBI allows for a more in-depth analysis of the changes in the firm size distribution of employment by enabling the regional dimension to be assessed and

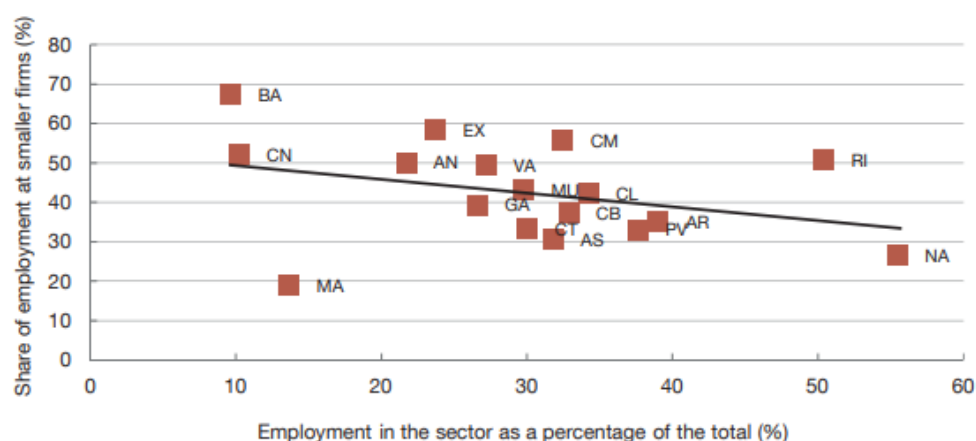
by isolating the effects of the different sectoral compositions across regions. Table 5 shows, for the 17 regions and four sectors under review, the change between the average shares of employment at smaller enterprises for the periods 2009-2014 and 2015-2021. By sector, the largest declines in the share of employment at smaller enterprises were in other services. These declines were particularly steep in Cantabria and Murcia. By contrast, in tourism-related services, the share of employment at smaller enterprises increased slightly in the country as a whole, albeit with a high level of dispersion across regions.

Employment at smaller enterprises decreased across all regions in the period under review, although it has tended to decline at a faster pace in the regions with a higher share in 2009 (see Chart 5). For example, in Castile-La Mancha and Extremadura, that share, which stood at around 75% in 2009, fell by 13 pp in both regions; all four sectors analysed contributed to this fall. At the other end of the scale, in the Basque Country and Madrid, this ratio fell very moderately (around 2 pp), given that the share of employment at smaller enterprises was already relatively low in 2009. In sum, this empirical evidence suggests that all regions appear to be trending towards a lower concentration of employment at smaller enterprises.

In short, the proportion of employees at smaller enterprises is considerably different across regions, even within the main sectors of activity. Madrid, followed by the Basque Country, Navarre and Catalonia, has the lowest proportion of employment at smaller enterprises, while at the other end of the scale, employment tends to be relatively more concentrated at such enterprises in Extremadura and Castile-La Mancha.

*Chart 4. Relationship between the sectoral specialisation of each region and the share of employment at smaller firms (fewer than 50 employees) in each sector (a)*

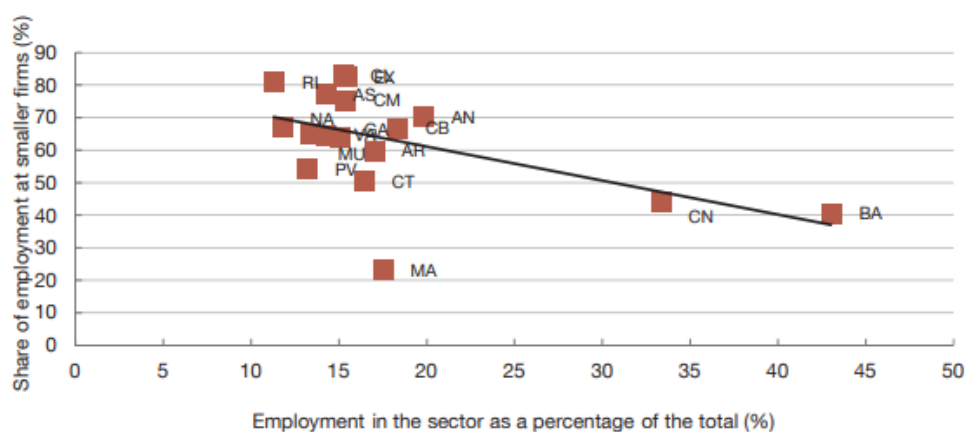
#### 4.a Industry



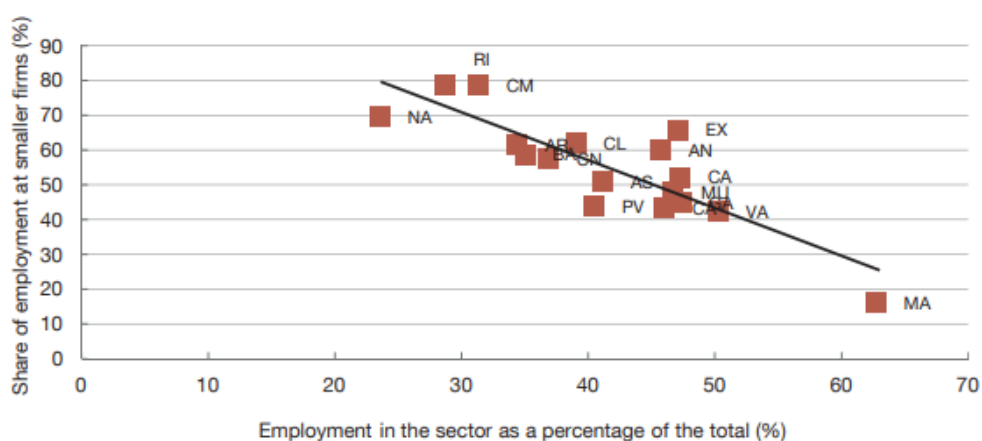
#### 4.b Construction



#### 4.c Tourism-related services



#### 4.d Other services

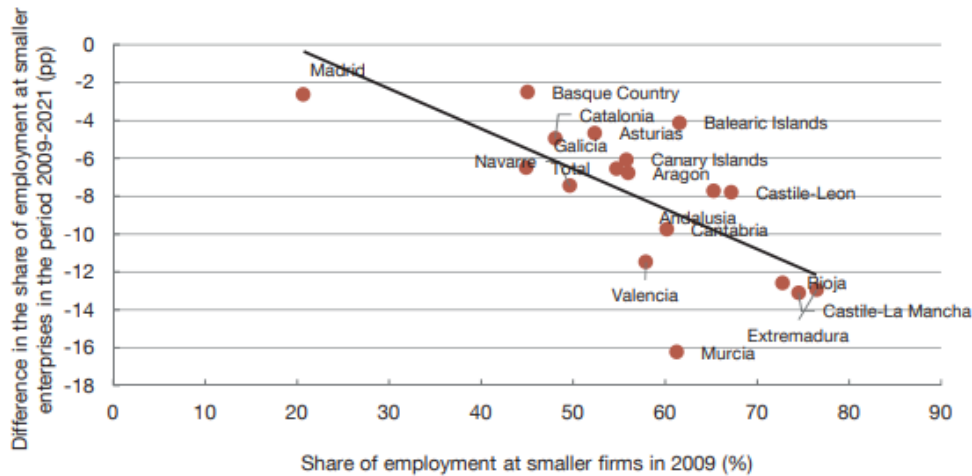


SOURCE: Banco de España.

(a) Average for the period 2009-2021. "Industry" includes the firms under NACE Rev. 2 sections B (mining and quarrying), C (manufacturing), D (electricity, gas, steam and air-conditioning supply) and E (water supply, sewerage, waste management and remediation). "Construction" includes firms in sector F (construction). "Tourism-related services" includes firms in sectors H (transportation), I (accommodation and food service).

activities) and R (arts, entertainment and recreation). "Other services" includes firms in sectors G (wholesale and retail trade), J (information and communication), M (professional, scientific and technical activities) and N (administrative and support service activities). AN: Andalusia, AR: Aragon, AS: Asturias, BA: Balearic Islands, CN: Canary Islands, CB: Cantabria, CL: Castile-Leon, CM: Castile-La Mancha, CT: Catalonia, VA: Valencia, EX: Extremadura, GA: Galicia, MA: Madrid, MU: Murcia, NA: Navarre, PV: Basque Country, RI: Rioja.

Chart 5. Change in share of employment at smaller enterprises (1-49 employees) in the CBI (a)



SOURCE: Banco de España.

(a) Including non-financial and non-agricultural market economy firms, excluding firms without any employees.

Table 5. Change in the percentage of employment at smaller firms (1-49 employees) in the CBI between 2009-2014 and 2015-2021 (pp) (a)

Region	Industry	Construction	Tourism-related services	Other services	Total
Andalusia	-3.7	0.9	-2.0	-14.2	-7.7
Aragon	-7.5	-8.8	-14.5	2.0	-6.8
Asturias	1.1	1.8	-5.6	-12.5	-4.7
Balearic Islands	1.7	-3.0	-7.3	-7.4	-4.1
Canary Islands	-2.6	6.8	-2.2	-11.3	-6.1
Cantabria	-5.1	8.2	-2.9	-22.2	-9.7
Castile-Leon	-8.9	3.0	-6.5	-5.5	-7.8
Castile-La Mancha	-18.1	0.4	-17.3	-14.7	-13.1
Catalonia	-1.8	2.0	1.5	-9.5	-5.0
Valencia	-13.1	-0.2	-5.1	-11.5	-11.5
Extremadura	-13.6	6.2	-7.5	-19.7	-12.9
Galicia	-6.9	-4.3	-0.7	-6.9	-6.6
Madrid	-5.4	4.8	2.4	-3.0	-2.6
Murcia	-11.0	-0.1	-9.0	-22.1	-16.2
Navarre	-5.6	4.6	-8.5	-6.5	-7.4
Basque Country	-1.3	0.1	-4.3	-2.7	-2.5
Rioja	-12.0	3.2	-17.2	-15.0	-12.6
Total	-5.8	1.0	0.3	-8.2	-6.5

SOURCE: Banco de España.

(a) Including non-financial and non-agricultural market economy firms, excluding firms without any employees

## 5. Conclusions

This paper shows that the CBI database is a useful source of information for analysing the structure of the Spanish business sector from a regional perspective. The analysis of the coverage rates and the comparison of the productive systems, employment dynamics and firm size distribution with other official sources allow us to conclude that the sample of firms in the CBI is representative of the population of NFCs at regional level. In addition, the empirical exercises presented in this paper indicate that there are considerable cross-regional differences in firm size distribution. This highlights the importance of factoring in the regional dimension when tackling the challenge of the small size of Spanish firms. Although the analysis has been conducted at regional level, the high number of firms available in the CBI and the broad coverage in the small, single-province regions point to the potential use of these firm-level data at provincial level.

Thus, the availability of these firm-level data that are representative at regional level for the Spanish economy opens the door to various lines of research. In particular, the financial and economic information contained in the CBI enables analysis of regional heterogeneity in different key areas. For example, the regional analysis of the determinants of business size and growth, productivity, profit margins and investment dynamics are some of the areas of interest that represent a priority line of work for the Banco de España.

## References

- [1] Almunia, Miguel, David López-Rodríguez and Enrique Moral-Benito. (2018). "Evaluating the Macro Representativeness of a Firm-Level Database: An Application for the Spanish Economy". Documentos Ocasionales, 1802, Banco de España. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3132982](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3132982)
- [2] Almunia, Miguel, Juan F. Jimeno, David López-Rodríguez and Borja Petit. (2024). "Size-dependent regulations in Spain". Mimeo.
- [3] Banco de España. (2022a). "Nota metodológica". In Banco de España, Resultados anuales de las empresas no financieras 2022. pp. 15-61. <https://www.bde.es/f/webbe/SES/Secciones/Publicaciones/PublicacionesAnuales/CentralBalances/22/Fich/ceba22nm.pdf>
- [4] Banco de España. (2022b). "Resultados anuales de las empresas no financieras 2022. Suplemento metodológico". <https://www.bde.es/f/webbe/SES/Secciones/Publicaciones/PublicacionesAnuales/CentralBalances/22/Fich/ceba22sm.pdf>
- [5] Banco de España. (2024). "Chapter 2. Structural challenges facing the Spanish economy". In Banco de España, Annual Report 2023, pp. 91-159. <https://repositorio.bde.es/handle/123456789/36537>
- [6] Fernández Cerezo, Alejandro, Matías Pacce and Isabel Sánchez. (2024). "La heterogeneidad regional en la evolución reciente de la inflación en España". Boletín Económico - Banco de España, 2024/T3, 08. <https://doi.org/10.53479/37592>
- [7] Gavilán, Ángel, and Domingo García Coto. (2024). "El reto del tamaño empresarial en España". Documento de Trabajo, 35, Fundación del Instituto Español de Analistas. <https://institutodeanalistas.com/wp-content/uploads/DOC.-No-35-El-reto-del-tamano-empresarial-en-Espana.pdf>
- [8] Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych and Sevcan Yesiltas. (2015). "How to Construct Nationally Representative Firm Level Data from the ORBIS Global Database". NBER Working Papers, 21558, National Bureau of Economic Research. <https://doi.org/10.3386/w21558>
- [9] Maza, Luis Ángel. (2024). "Una reflexión sobre los umbrales cuantitativos en los modelos de depósito de las cuentas anuales y su posible impacto en el tamaño empresarial en España". Documentos Ocasionales, 2419, Banco de España. <https://doi.org/10.53479/36639>
- [10] Mora-Sanguinetti, Juan S., and Isabel Soler. (2022). "La regulación sectorial en España. Resultados cuantitativos". Documentos de Trabajo, 2202, Banco de España. <http://dx.doi.org/10.2139/ssrn.4035629>
- [11] Moral-Benito, Enrique. (2018). "Growing by learning: firm-level evidence on the size-productivity nexus". SERIEs, 9(1), pp. 65-90. <https://doi.org/10.1007/s13209-018-0176-2>
- [12] Prades Illanes, Elvira, and Patrocinio Tello Casas. (2020). "The heterogeneous economic impact of COVID-19 among euro area regions and countries". Economic Bulletin - Banco de España, 2/2020, Analytical Articles. <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/ArticulosAnaliticos/20/T2/descargar/Files/be2002-art17e.pdf>



# THE REGIONAL PERSPECTIVE OF THE BANCO DE ESPAÑA'S CENTRAL BALANCE SHEET DATA OFFICE (CBSO)

Alejandro Fernández Cerezo  
Borja Fernández-Rosillo San Isidro  
Natividad Pérez Martín

**WORKSHOP: "NEW INSIGHTS FROM FINANCIAL STATEMENTS"**

BANCO DE ESPAÑA  
Madrid, 17th october 2024





# CONTENT

1. Introduction
2. “BExplore Regional Central Balance Sheet Data Office”
3. Regional coverage and representativeness of the CBI database
4. Exploring heterogeneity in firm size distribution at regional level
5. Conclusions

Why is it important to have a representative firm-level database at regional level?

- Firm-level databases enable a deeper understanding of economic developments, also at regional level:
  - Regional differences in corporate variables may reveal trends not observable at national level.
  - Shocks affecting the whole country (pandemic/energy) may have asymmetric impacts across regions.
  - In decentralized countries, different levels of regulation across autonomous communities (CCAA).
- Almunia, López-Rodríguez and Moral-Benito (2018) evaluated the macro-representativeness of the Banco de España's Integrated Central Balance Sheet Database (CBI, in Spanish).
  - The CBI was able to replicate, at national level, the growth rate of employment in the Spanish private sector and the firm size distribution.
- In Fernández-Cerezo, Fernández-Rosillo and Pérez-Martín (2024), we show that the CBI database is representative also at regional (CCAA) level:
  - The CBI exhibits high coverage of firms with respect to the population of firms in each CCAA.
  - The CBI replicates the sectoral distribution of employment across CCAA.
  - The CBI replicates the dynamics of aggregate employment at CCAA level.

# 1. INTRODUCTION

## The Banco de España's CBI database



### What is the CBI\* database?

Information based on financial accounts of Spanish NFCs: balance sheet, income statement, # employees, **location\*\***...



### Sources of information of CBI

Voluntary questionnaire directly from firms +  
Mandatory deposits of annual accounts in Mercantile Registries



### Data treatment & cleansing to raw firm data

From 1M NFCs received per year to 840k firms suitable for analysis after data cleansing by CBSO staff



### Main characteristics of the CBI

Unit of analysis: individual enterprise (identified by anonymised tax id)  
Frequency: annual, from 1995 to 2022



### How can CBI data be accessed and visualized?

For researchers: microdata through [BELab data laboratory](#)  
For general public: aggregate data through ["BExplore" visualizations](#)



\* CBI: Integrated Central Balance Sheet Database

\*\* Location: each firm is located in the province where it is headquartered.

## 2. “BEXPLORE REGIONAL CENTRAL BALANCE SHEET DATA OFFICE”

An interactive website to provide the general public with easy access to aggregate CBI regional data

In order to bring regional CBI analysis closer to users, an interactive tool named **“BExplore Central Balance Sheet Data Office Regional”** has been developed

### MAIN CONTENT

Available statistical information:

Nº of firms (% coverage)

Weighted Average Value

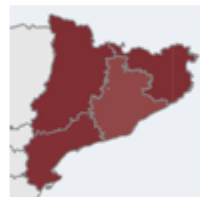
Median value (Q2)

about:



17 autonomous communities (CCAA) + Ceuta/Melilla

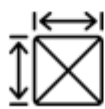
for:



52 provinces



14 sectors of activity



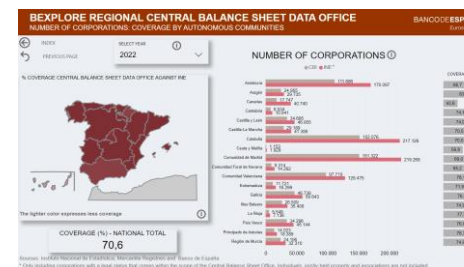
4 company sizes



15 years (2008-2022)

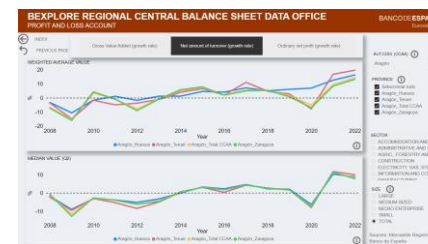


11 variables



Coverage rates

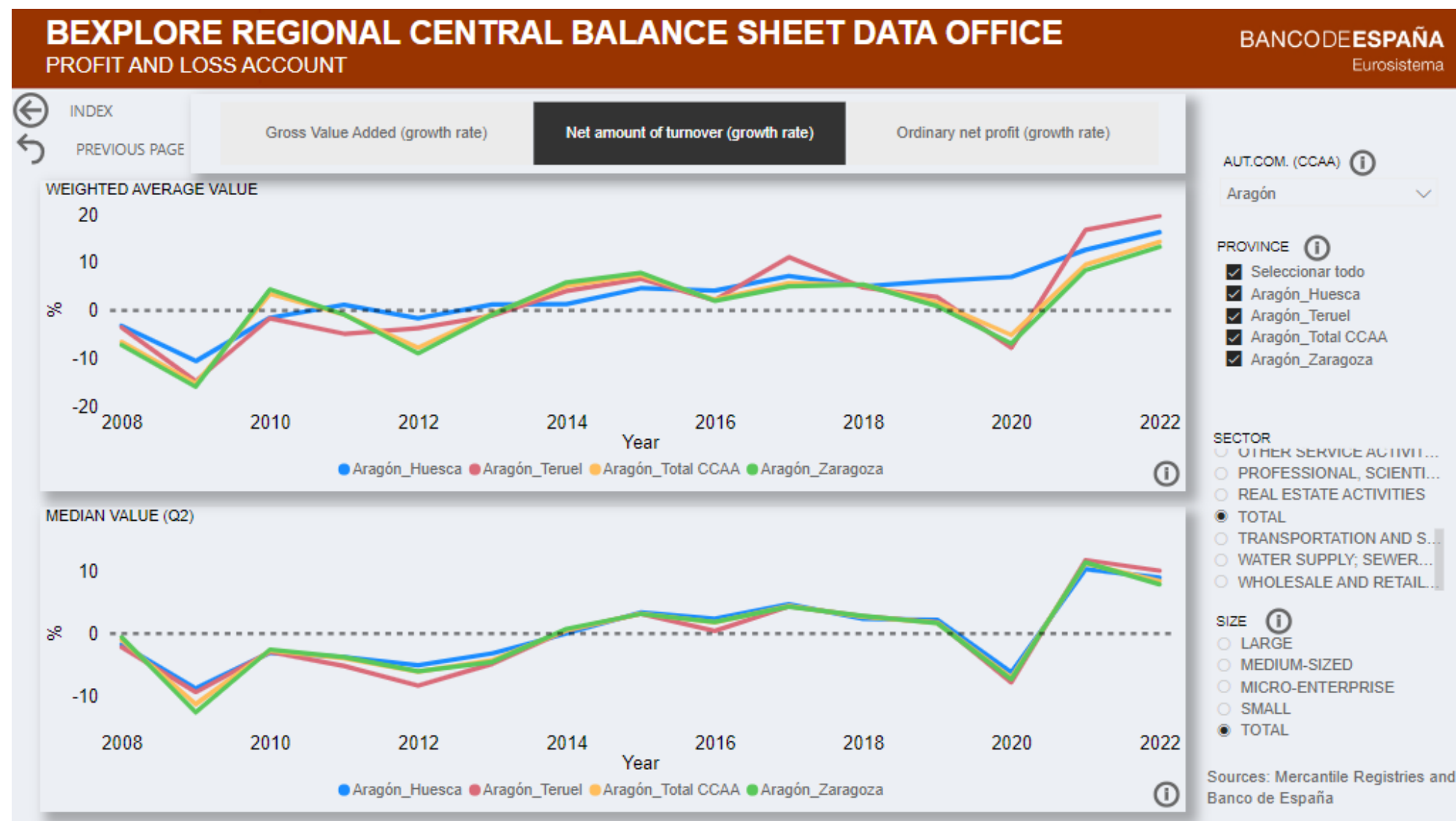
Sectoral breakdown



-P&L variables  
-Employment and labor costs variables  
-Financial ratios: margins, financial cost, debt ratio...

## 2. “BEXPLORE REGIONAL CENTRAL BALANCE SHEET DATA OFFICE”

Example



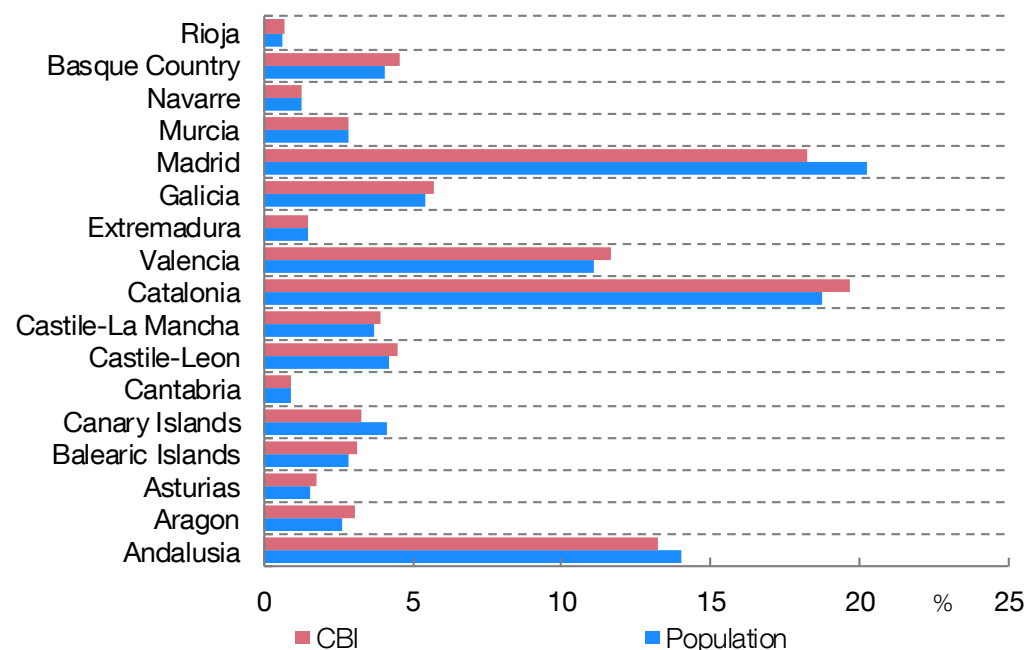
### 3. REGIONAL COVERAGE AND REPRESENTATIVENESS OF THE CBI DATABASE

The CBI database is reasonably representative of the population of Spanish NFCs at regional level

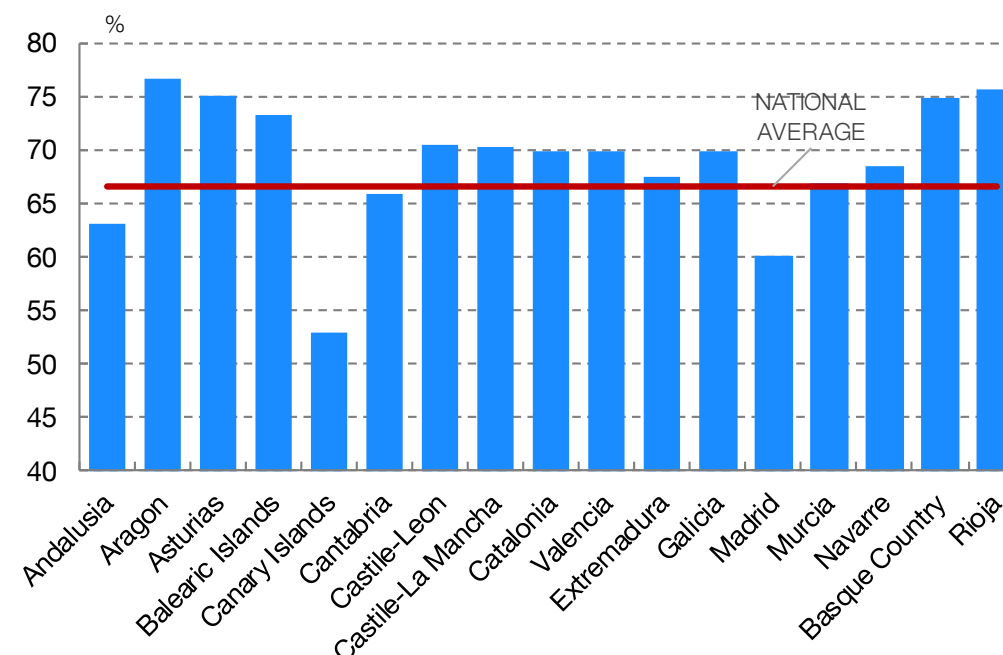
The regional distribution of the # of Spanish NFCs in the CBI adequately replicates that of the population

The coverage rate at national level is 66.5% on average, exceeding 70% in several CCAA

**DISTRIBUTION OF THE NUMBER OF NFCs IN CBI AND IN POPULATION BY CCAA (a)**



**COVERAGE RATES BY CCAA  
(# NFCs IN CBI / # NFCs IN POPULATION) (a)**



Sources: Banco de España and INE

(a): Average for the period 2009-2021. The number of firms comprising the population of NFCs in Spain comes from the Central Business Register, prepared by the National Statistics Institute (INE), which in turn feeds Eurostat's structural business statistics.



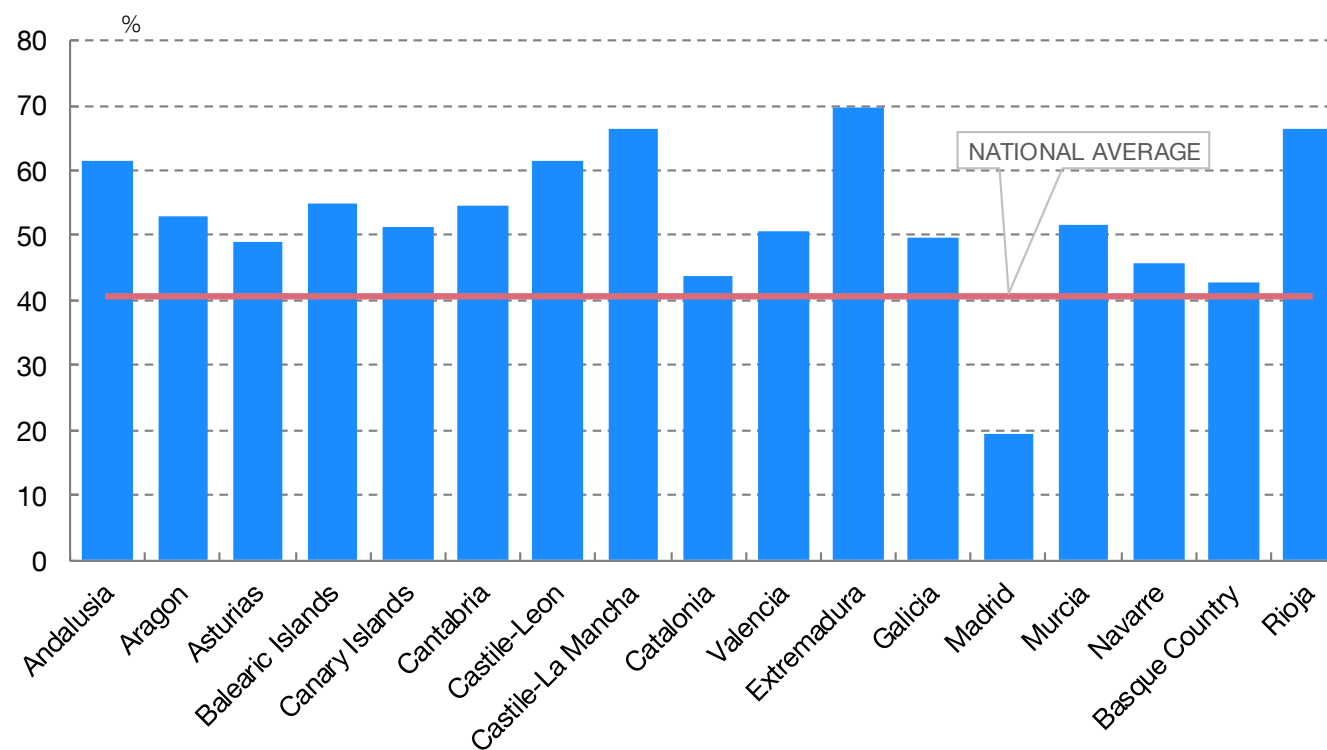


## 4. EXPLORING HETEROGENEITY IN FIRM SIZE DISTRIBUTION AT REGIONAL LEVEL

The CBI is a unique source in Spain for analysing regional heterogeneity in firm size distribution

- According to CBI, small firms (<50 employees) account for 70% of employment in Extremadura and only 20% in Madrid in the non-financial and non-agricultural market economy in Spain (vs 40% in the national average).

SHARE OF EMPLOYMENT AT SMALLER FIRMS (1-49 EMPLOYEES) IN CBI (a)



Source: Banco de España.

(a): Average for the period 2009-2021. Including non-financial and non-agricultural market economy firms, excluding firms without any employees.

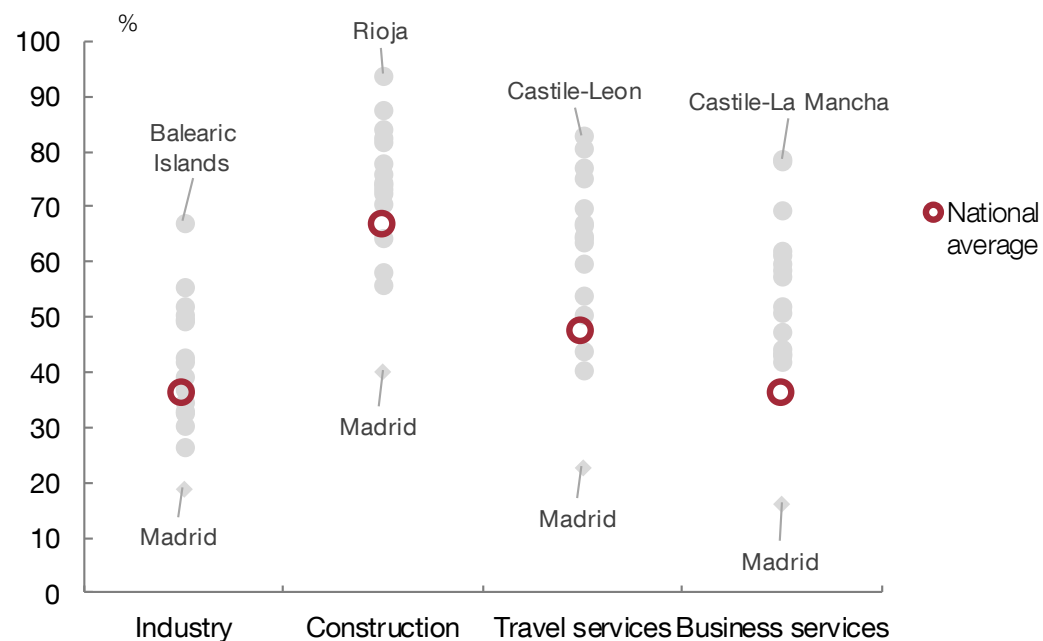


## 4. EXPLORING HETEROGENEITY IN FIRM SIZE DISTRIBUTION AT REGIONAL LEVEL

CBI provides a breakdown of employment by firm size, sector and region

- Within each sector there are notable cross-regional differences in firm size.

**SHARE OF EMPLOYMENT AT SMALLER FIRMS  
(<50 EMPLOYEES) IN CBI (a)**



Source: Banco de España.

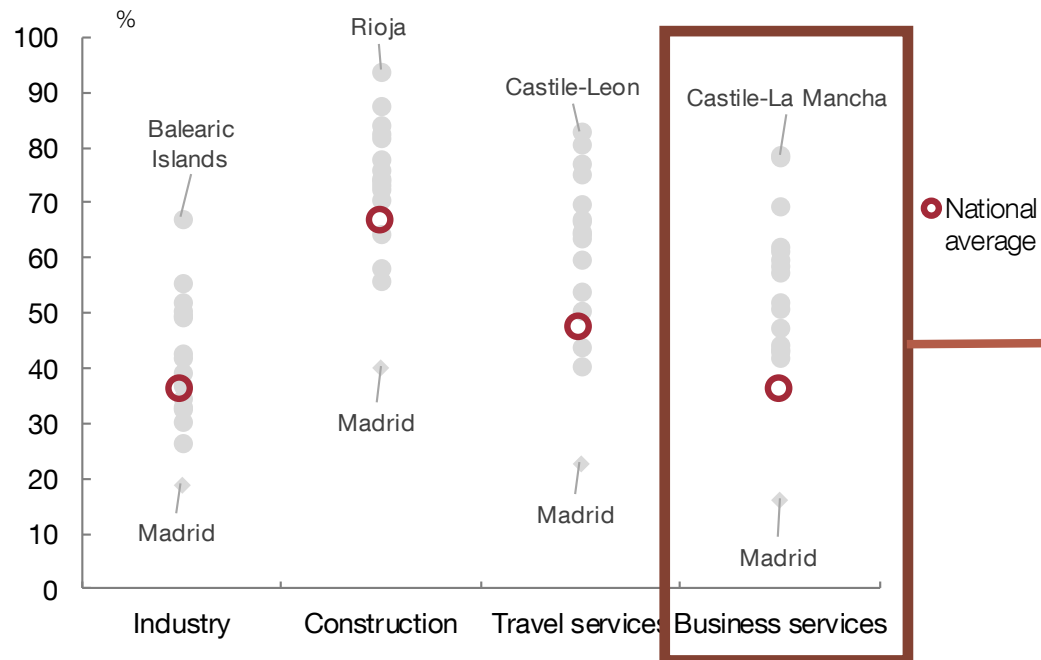
(a): Average for the period 2009-2021. Including non-financial and non-agricultural market economy firms, excluding firms without any employees. “Industry” includes the firms under NACE Rev. 2 sections B, C, D and E. “Construction” includes firms in F. “Travel services” includes firms in H, I and R. “Business services” includes firms in G, J, M and N. AN: Andalusia, AR: Aragon, AS: Asturias, BA: Balearic Islands, CN: Canary Islands, CB: Cantabria, CL: Castile-Leon, CM: Castile-La Mancha, CT: Catalonia, VA: Valencia, EX: Extremadura, GA: Galicia, MA: Madrid, MU: Murcia, NA: Navarre, PV: Basque Country, RI: Rioja.

## 4. EXPLORING HETEROGENEITY IN FIRM SIZE DISTRIBUTION AT REGIONAL LEVEL

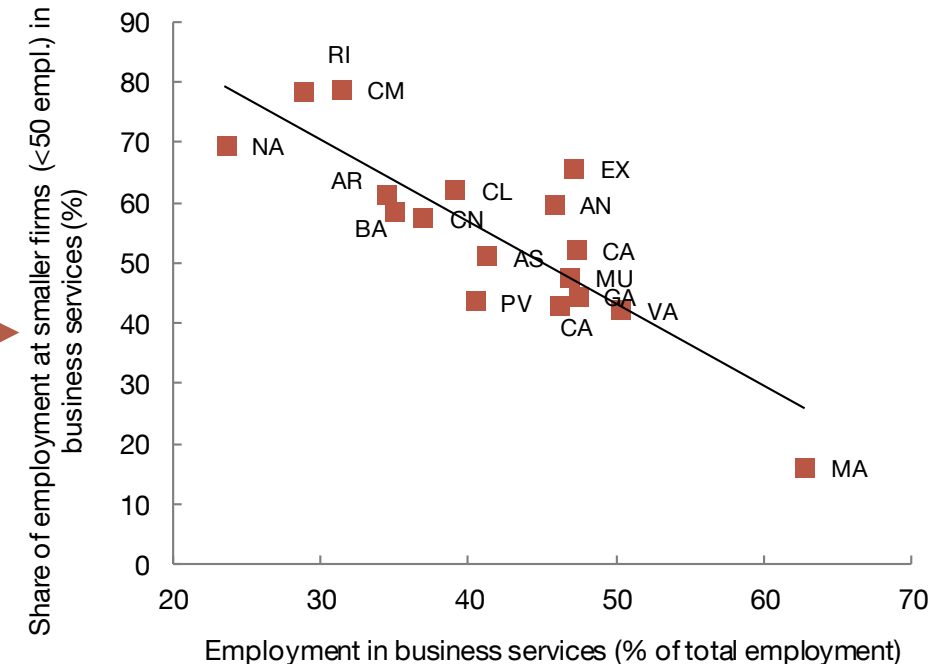
CBI provides a breakdown of employment by firm size, sector and region

- Within each sector there are notable cross-regional differences in firm size.
- The share of employment at smaller enterprises tends to be lower in those sectors in which the region is more specialized.

**SHARE OF EMPLOYMENT AT SMALLER FIRMS  
(<50 EMPLOYEES) IN CBI (a)**



**SECTORAL SPECIALISATION AND THE SHARE OF  
EMPLOYMENT AT SMALLER FIRMS IN THE BUSINESS  
SERVICES SECTOR (a)**



Source: Banco de España.

(a): Average for the period 2009-2021. Including non-financial and non-agricultural market economy firms, excluding firms without any employees. "Industry" includes the firms under NACE Rev. 2 sections B, C, D and E. "Construction" includes firms in F. "Travel services" includes firms in H, I and R. "Business services" includes firms in G, J, M and N. AN: Andalusia, AR: Aragon, AS: Asturias, BA: Balearic Islands, CN: Canary Islands, CB: Cantabria, CL: Castile-Leon, CM: Castile-La Mancha, CT: Catalonia, VA: Valencia, EX: Extremadura, GA: Galicia, MA: Madrid, MU: Murcia, NA: Navarre, PV: Basque Country, RI: Rioja.



### **Representative database**

- The CBI database is a representative source of information for analyzing the structure of the Spanish business sector from a regional perspective.



### **Interactive tools**

- In an effort to ease both external researchers' and the general public's access to Bank of Spain's databases, we are developing new user-friendly, interactive tools to present the potential uses of the CBI databases.



### **Cross-regional differences**

- We showed that there are considerable cross-regional differences in firm size distribution.



### **Diverse potential research lines**

- The availability of CBI firm-level data opens the door to various lines of research for analysis of regional heterogeneity in different key areas: e.g. the determinants of business size and growth, productivity, profit margins and investment dynamics.

THANKS FOR YOUR ATTENTION

