

12th biennial IFC Conference: “Statistics and beyond: new data for decision making in central banks”

22-23 August 2024

Addressing sustainability data gaps for SMEs: challenges, alternative solutions and data-sharing¹

Borja Fernández-Rosillo San Isidro,
Bank of Spain

¹ This contribution was prepared for the conference. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.

Addressing sustainability data gaps for SMEs

Challenges, alternative solutions and data-sharing

Borja Fernández-Rosillo San Isidro

Abstract

The primary aim of this paper is to provide an overview of the current state of sustainability indicators for SMEs, the challenges ahead and possible alternative approaches and data-sharing solutions to address it evaluated and not yet implemented in Spain. Despite SMEs not having direct regulatory mandates to prepare ESG information, they are increasingly compelled to do so due to indirect factors. These include their close relations with financial institutions, their desire to access financial services, their interactions within larger business networks and value chains, and their commitment to maintaining competitiveness, profitability, and market transparency. However, SMEs face significant challenges in terms of technical capacity and resource constraints when adapting to these new requirements. This paper presents some conceivable approaches explored to address these data gaps from a Central Balance Sheet Data Office perspective, focusing on alternative solutions and data-sharing among institutions:

In terms of alternative approaches, the following solutions have been explored:

- (i) An econometric model for estimating the carbon footprint of non-financial corporations, company by company, using Green House Gas (GHG) emissions by NACE sector estimated in Environmental Accounts, in combination with a series of accounting variables (i.e. average age of fixed assets) that presumably are relevant to explain the carbon footprint of a company;
- (ii) The inclusion of a set of voluntary sustainability-related indicators in the annual accounts deposit in the Mercantile Registers by SMEs.
- (iii) An estimation derived from the utilization of specific accounting items from the P&L account, such as electricity and fuel expenses categorized under Operating Expenditures (OPEX), or fuel consumption included in the cost of goods.

In terms of data-sharing, the following solutions have been explored:

- (i) A public-private data collaboration for the creation of a Green Data Space for data sharing among different public and private institutions.
- (ii) An initiative involving key Spanish public institutions with the objective of establishing an integrated administrative register of pollutants and GHG emissions.

The need to close the sustainability data gaps among SMEs is crucial for enhancing the availability of information and improve decision making. This step will enable stakeholders, regulators, and markets to assess how SMEs are managing sustainability risks and opportunities. Furthermore, leveraging the existing data gap to encourage data-sharing across various registers and public-private institutions is essential for promoting better and more comprehensive information.

Keywords: ESG (Environmental, Social, Governance), data gaps, standards, climate change, sustainability, databases, metrics, digitalized, regulation, SMEs, data-sharing.

JEL classification: Q5 (Environmental economics).

Index

Challenges, alternative solutions and data-sharing.....	1
1. Introduction	4
2. Current situation of ESG indicators for SMEs	4
3. Stakeholders needs and main drivers to address them	8
4. Alternative measures from a Central Balance sheet perspective to fill the data gaps on SMEs ESG.....	9
4.1 An econometric model for estimating the carbon footprint of non-financial corporations.....	9
4.2 Voluntary sustainability-related indicators in the annual accounts deposit in the Mercantile Registers.....	12
4.3 Estimation methodology of energy consumption and GHG emissions based on specific items from the P&L account	13
4.4 Data-sharing: Public-private collaboration for the creation of a Green Data Space.....	15
4.5 Data sharing: the relevance of establishing an integrated repository of pollutants and GHG emissions among key Spanish public institutions.....	16
5. Conclusions.....	19
References	20

1. Introduction

Social concern and risks regarding ESG (Environmental, Social and Governance) issues have significantly increased in recent years. This is why numerous regulatory authorities have made an additional effort to try to regulate and require companies to produce this type of information in order to reduce negative impacts on the environment, guarantee a sustainable transition of the economy towards greener practices and ensure respect for human rights and ethics in business.

The problem is that the vast majority of these requirements and information needs have been oriented towards large companies and listed groups on regulated markets. The principle of proportionality (not to impose additional information burden to small businesses) has generated a lack of obligations and information disposal for SMEs (Small and Medium Enterprises), which usually represent an important driver of a country's economy and a boost for employment and Gross Domestic Product (GDP).

This ESG regulatory orientation towards big and listed companies responds to the markets necessities and the relevance of the impact of this companies in sustainability related areas. SMEs, on the contrary, have less significance from an individual perspective but from an aggregated point of view the impact and influence on ESG topics are notorious. Additionally, their role in the value chain of large corporations should not be overlooked. Consequently, there have been some advances in regulatory requirements for non-listed SMEs (mostly from a voluntary perspective to avoid excessive burdening) to disclose ESG information¹. However, these new requirements are voluntary, not currently available and limited in comparison to larger company's regulatory requirements. For this reasons, alternative solutions need to be addressed to fulfil this data gap.

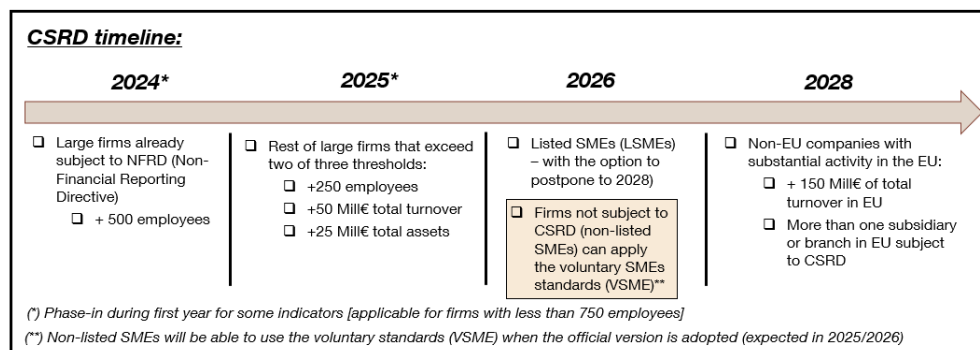
The objective of this paper is to present an overall picture in terms of non-listed SMEs ESG data availability and possible alternative solutions from a Central Balance sheet perspective to address this data gaps.

2. Current situation of ESG indicators for SMEs

In relation to ESG disclosures there have been some considerable advances, especially in Europe with the CSRD (Corporate Sustainability Reporting Directive)² that introduces mandatory ESG requirements for big companies and listed SMEs in a standardized and digitalized format (the entry into force of the CSRD will be progressive from 2024 onwards with the first reports available on 2025). Additionally, it includes the possibility for non-listed SMEs to report in accordance with a list of voluntary proportionate standards (known as VSME).

¹ For listed SMEs there is a regulatory obligation in Europe to present ESG information

² Corporate Sustainability Reporting Directive - [link](#)



Graph 1. CSRD timeline

For the moment, the main standards and recommendations developed so far for preparing and presenting sustainability information are:

- European Sustainability Reporting Standards (ESRS) developed by EFRAG under the mandate of the CSRD.
- ISSB (International Sustainability Standards Board) Sustainability standards S1 and S2 developed under IFRS which apply globally.
- Global Reporting Initiative (GRI) standards which apply globally.
- SASB (Sustainability Accounting Standards Board) standards which apply globally.
- TCFD (Task Force on Climate-Related Financial Disclosures) standards and recommendations which apply globally.

The main issue is that none of these standards have an SME version except for the ones developed by EFRAG which recently have published a specific version of their sustainability standards focused on SMEs (VSME)³. In the case of the GRI they have published two documents⁴ highlighting some best practices and the importance of ESG reporting for SMEs but no specific standards for SME have been developed.

The VSME EFRAG standard (currently just a draft pending to be finalized during 2024) represents a simplified and SME oriented version based on the ESRS developed for big companies and listed SMEs subject to the CSRD. The rest of international frameworks, although they could be applied by any type of corporation, they are mostly oriented to bigger and listed companies and lack any simplified version for SME as the one published by EFRAG.

The VSME covers the same sustainability topics as the ESRS developed for large companies but in a simplified and proportionate manner. The draft standards contain three modules:

- **Basic Module (BM):** it is a minimum requirement for all companies using the VSME. It includes information related to environmental (energy, GHG, pollution, biodiversity, water, resources, circular

³ EFRAG published Exposure Drafts (ED) for LSME and VSME on 22 January 2024. Currently they are under a revision phase after the public consultation in order to have the final versions publicly available in 2025 → link

⁴ For further information, please refer to GRI documents published on their website related to: "Small business big impact: SME sustainability reporting from vision to action" and "Empowering Small Business"

economy...), social (information on company's workforce) and business conduct items (corruption and bribery).

- **Narrative – Policies, Actions and Targets (PAT) Module:** It includes narrative disclosures to be reported in addition to the basic module related to policies, actions and targets the company is conducting over sustainability topics.
- **Business Partners (BP) Module:** it contains additional datapoints to be reported in terms of information requested from business partners such as investors, lenders, corporate clients, suppliers...

This standard does not respond to a regulatory obligation and is applied on a voluntary basis and, consequently, different options are available which consist of: (i) just presenting the Basic Module; (ii) presenting the Basic Module + PAT Module; (iii) presenting the Basic Module + BP Module; (iv) presenting all modules. This makes information reporting more efficient and proportionate in order to facilitate SMEs their participation in the transition towards a more sustainable economy.

In table 1 a list of the different disclosure requirements contained in the VSME is presented.

Module	Category	Disclosure
BASIC MODULE (BM)	General disclosures	Disclosure B 1 – Basis for Preparation
		Disclosure B 2 – Practices for transitioning towards a more sustainable economy
	Basic metrics: Environment	B 3 – Energy and greenhouse gas emissions
		B 4 – Pollution of air, water and soil
		B 5 – Biodiversity
		B 6 – Water
		B 7 – Resource use, circular economy, and waste management
	Basic metrics: Social matters	B 8 – Workforce – General characteristics
		B 9 – Workforce - Health and Safety
		B 10 – Workforce – Remuneration, collective bargaining, and training
	Basic metrics: Business conduct	B 11 – Convictions and fines for corruption and bribery
Narrative – Policies, Actions and Targets (PAT) Module		Disclosure N 1 – Strategy: business model and sustainability related initiatives
		Disclosure N 2 – Material sustainability matters
		Disclosure N 3 – Management of material sustainability matters
		Disclosure N 4 – Key stakeholders
		Disclosure N 5 – Governance: responsibilities in relation to sustainability matters
Business Partners (BP) Module		Disclosure BP 1 – Revenues from certain sectors
		Disclosure BP 2 – Gender diversity ratio in governance body
		Disclosure BP 3 – GHG emissions reduction target
		Disclosure BP 4 – Transition plan for climate change mitigation
		Disclosure BP 5 –Physical Risks from climate change
		Disclosure BP 6 – Hazardous waste and/or radioactive waste ratio
		Disclosure BP 7 – Alignment with internationally recognized instruments

Module	Category	Disclosure
		Disclosure BP 8 – Processes to monitor compliance and mechanisms to address violations
		Disclosure BP 9 – Violations of OECD Guidelines for Multinational Enterprises or the UN Guiding Principles (including the principles and rights set out in the 8 fundamental conventions of the ILO Declaration and the International Bill of Human Rights)
		Disclosure BP 10 – Work-life balance
		Disclosure BP 11 – Number of apprentices

Table 1. List of voluntary disclosure requirements contained in EFRAG VSME

Notice that these standards include the main relevant ESG datapoints useful for a whole range of stakeholders. Some examples of specific datapoints incorporated within these standards are:

- Total energy consumption in MWh with a breakdown between fossil fuels and electricity;
- Scope 1 and Scope 2 emissions in tCO₂eq;
- Pollutant emitted in the undertakings own operations;
- Impacts on biodiversity and the ecosystem;
- Total water withdrawal, consumption and location in high water stress areas;
- Total annual waste generation by type (hazardous and non-hazardous);
- Total number of employees by employment contract, gender or country;
- Total number of work-related accidents;
- The percentage gap in pay between genders;
- Total amount of fines for violation or anti-bribery laws;
- ...

These SME oriented standards are an important step towards increasing availability and comparability of SME sustainability data and a way to enable them face the increasing request for ESG data and lower their entry barrier to reporting. However, there is still a long path to go as:

- Other frameworks do not currently have any specific SME oriented standards. The VSME should be taken as an example for other international standards and jurisdictions to bring the sustainability reporting closer to SMEs bearing in mind their characteristics, business relations and limited resources in comparison with bigger companies.
- Most SMEs lack technical capacity and resources to meet with quality and completeness all these requirements.
- In the majority of jurisdictions there is not a direct obligatory requirement for SMEs to prepare and report this information.

Meanwhile alternative approaches and solutions need to be addressed and explored. In section three a series of indirect drivers will be presented as a possible catalyst to ease SMEs to report this information and, in section four, some alternative approaches and proxies to fulfill this data gap will be presented.

3. Stakeholders needs and main drivers to address them

When analysing sustainability data availability, one of the main issues is the population coverage, in which SMEs (Small and Medium Enterprises) play an important role as they normally account for more than 95% of a country's firms. In Spain they represent more than 99% of total firms as shown in Table 2.

Year 2022	By n° of employees	N° of firms	In %
SMEs	Without employees	429,133	33.2%
	1 – 9 employees	719,776	55.7%
	10 – 49 employees	119,479	9.2%
	50 – 249 employees	19,302	1.5%
Large firms	250 – 449 employees	2,205	0.2%
	500 – 999 employees	1,072	0.1%
	+ 1,000 employees	766	0.1%
TOTAL		1,291,733	100%

Table 2. Number of companies in Spain by size in 2022 – Central Business Register (CBR) - INE

Most jurisdictions exempt SMEs from presenting sustainability information due to their limited technical and organizational capacity to produce it in comparison to bigger firms which have more resources and experience in the field. This is an important data gap to bear in mind as SMEs normally represent most of the undertakings of a country, and additionally, although individually their impact is limited, collectively their influence and repercussion is considerable. However, this situation is slowly changing due to the following reasons provoked by the different needs of a wide spectrum of stakeholders:

- **Access to financing.** SMEs will be forced indirectly to produce this information as financial institutions (banks) or investors will require it in order to comply with regulatory obligations. Moreover, banks and investors are incorporating ESG risks in their internal models so they need this data to correctly analyse the risk profile of its SME clients or investments in order to see if the financial operation is viable or not under their current risk standards (i.e. a loan, an investment...). Additionally, transition and physical risk are increasingly being taken into consideration by banks/financial institutions and SMEs may need to correctly report information related to those risks if they want to continue to have access to financial institutions.
- **Value chain requirements from bigger groups/companies.** Bigger groups need information relative to SMEs to fulfill other regulatory requirements, such as the CSRD or the CSDDD (Corporate Sustainability Due Dilligence Directive), in order to operate with companies that are environmentally responsible and comply with ESG standards. This pressure would force SMEs to focus on the preparation of this type of information. However, there is a limitation to what theoretically a bigger group/company can ask an SME as the recently developed regulation (CSRD) establishes that sustainability reporting obtained and asked to SMEs by third parties should not exceed the information to be disclosed in accordance with the voluntarily sustainability reporting standards for small and medium-sized undertakings (VSME). This

approach is logical and limits what big companies can ask an SME but there is a minimum of information that SMEs would have to prepare and proportionate in relation to sustainability topics if they want to continue their business relationship with bigger companies and be competitive in the market.

- **Market transparency and good practices.** SMEs should provide clear and transparent information on their ESG risks and performance so that market participants and stakeholders can correctly evaluate them before establishing any type of business relationship. The actual globalization environment has increased competition and broad new challenges that SMEs are exposed to and must face in order to survive and be profitable. In Europe, the voluntarily SME standards for the preparation of this information will increase uniformity, comparability, and reliability of data. This progress will support the competitive growth of SMEs and increase their resilience in the medium-long-term.

In conclusion, the lack of environmental data for SMEs is a problem to be addressed. The need for this information is of particular interest to a Central Balance Sheet Data Office (CBSO) whose microdata serves as a basis for economic analysis of non-financial companies. Hopefully, the growing indirect requirements coming from the value chain, financial institutions and investors needs (market transparency) will invite and ease SMEs to calculate and report environmental data. However, in the meantime, this data gap must be dealt by using proxies and estimates. Some alternative approaches from a CBSO perspective are presented in section four.

4. Alternative measures from a Central Balance sheet perspective to fill the data gaps on SMEs ESG

4.1 An econometric model for estimating the carbon footprint of non-financial corporations

The objective of this project is to build an econometric model capable of estimating the carbon footprint of each Spanish non-financial company based on the greenhouse gas (GHG) emission intensity of its sector of activity in combination with other accounting variables. The goal is to modulate the generic GHG value associated to that sector by using the accounting variables of that concrete firm and estimate a customised carbon footprint at company level. To achieve this specification, some of the accounting variables of the Central Balance Sheet Data Office are considered.

The focus of this paper is not to describe in detail the model as it would require a paper in itself. However, the main considerations and specifications related to the experimental model developed are provided below to guide the reader on the content and characteristics of this model and to help and serve as inspiration for other users currently working or thinking of developing something similar.

The variables used in the model are presented in the table 3:

Type	Variables	Brief description
Response variable (yt)	Emission coefficient (TnCO ₂ eq per 1,000€ produced)	It is the ratio of: (i) Total GHG emissions in TnCO ₂ e by economic sector (based on NACE ⁵ classification) available in the environmental accounts published by the National Statistics Institute (NSI) (ii) divided by the total aggregated production ⁶ of its economic sector (NACE)
Explanatory variables (xt)	Average age of fixed assets in years	Ratio calculated internally for each individual non-financial company, by dividing the stock of fixed assets (year t) by the annual depreciation (year t).
	NACE at 2 digits	NACE at 2 digits (transformed to a dummy variable)
	Years	Years of study from 2015 to 2019 (transformed to a dummy variable)
	Size proportion	The proportion (%) of the number of companies that are Small, Medium and Large divided by the total amount of companies of that specific NACE at 2 digit level
	Average number of employees	The average number of employees
	Accounting variables	• Net amount of turnover (measured as a % over total production) • Depreciations (measured as a % over total production) • Sales (measured as a % over total production) • Intermediate consumptions (measured as a % over total production) • Tangible assets (measured as a % over total assets) • Intangible assets (measured as a % over total assets) • Other operating expenses (measured as a % over total production) • Personnel costs (measured as a % over total production) • Gross Value Added (measured as a % over total production) • Current assets (measured as a % over total assets) • Suppliers (measured as a % over total assets)

Table 3. Variables used for the econometric model

Data preparation was needed in terms of:

- Sector classification. Microdata aggregation and adaptation to the 2 digit NACE sector classification was conducted.
- Data normalization. Aggregated accounting variables were divided by the total asset (for balance sheet items) and by the production (for Profit and Loss items) in order for them to be expressed in the same magnitude.

Data showed:

- The heterogeneity of emissions among different sectors.
- The gradual decline in the Emission coefficient (measured in TnCO₂eq per 1,000€ produced) since 2015.

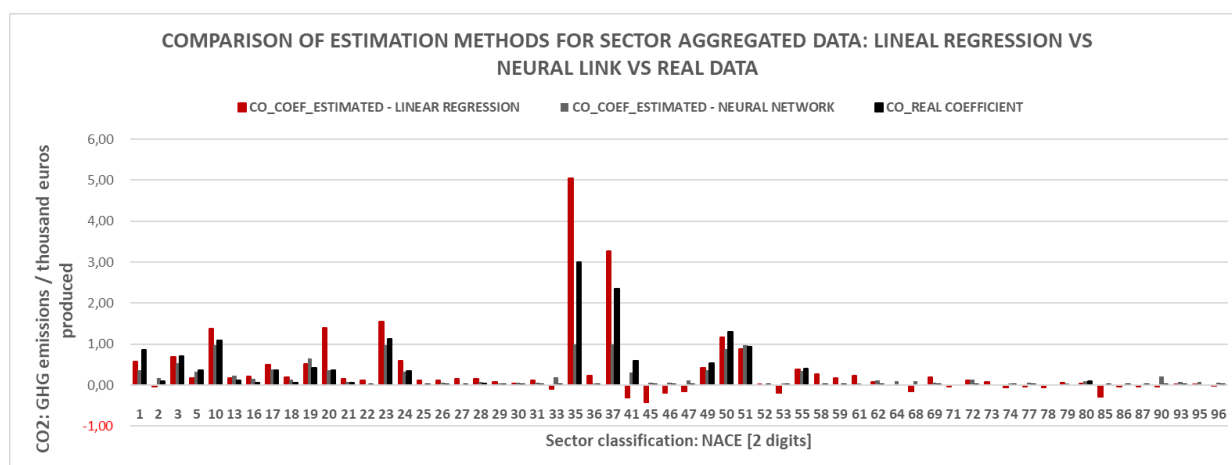
Different approximations and models were tried:

- Linear models:
 - Linear regression
 - Linear regression by steps
- Non-linear models:
 - Ridge regression
 - Lasso regression

⁵ NACE: National Classification of Economic Activities

⁶ Production = Net amount of turnover - Consumption of goods + Variation in finished and in-progress goods stocks level + Tasks performed by the company for assets + Other operating income unclassified

- Random Forest
- Neural network



Graph 1. Comparison of estimation methods used for sector aggregated data

In graph 1, several issues arise with the different methods. Linear regressions yield negative values that lack meaning in this context (having negative emissions has no sense). On the other hand, they overestimate sector estimates for the most polluting NACEs (NACE 35 - Electricity, gas, steam, and air conditioning supply, and NACE 37 - Sewerage, waste management, and remediation activities). On the contrary, the neural network does not have the issue of negative values but falls short in estimating the most polluting CNAEs and, additionally, it faces the common challenge of interpretability, often referred to as the “black box” problem (even in a small neural network like this one).

COMPANY	[A] CO_REAL COEFFICIENT	[B] CO_COEF_EST - LINEAR REGRESSION	[C] CO_COEF_EST - NEURAL NETWORK	Difference [B]/[A]	Difference [C]/[A]
COMPANY 1	0.11	0.57	0.08	416%	-27%
COMPANY 2	0.86	1.38	1.00	60%	16%
COMPANY 3	1.20	3.40	1.00	183%	-17%
COMPANY 4	0.02	0.61	0.98	2925%	4800%
COMPANY 5	0.66	4.89	1.00	641%	52%
COMPANY 6	13.84	1.39	1.00	-90%	-93%
COMPANY 7	0.26	0.34	0.30	29%	15%
COMPANY 8	0.13	-0.01 [negative coeff]	0.02	-106%	-85%
COMPANY 9	0.01	0.07	0.02	610%	100%
COMPANY 10	0.03	-0.11 [negative coeff]	0.02	-477%	-33%
...	...	...	...	...	...

Table 4. Some examples of individual comparison between estimated coefficients and real data

Moreover, when analysing and comparing the estimated coefficients with real data at company level (as shown in table 4) important differences arise in some cases hindering the use of this models as a proxy for estimating GHG at company level.

The main conclusions of this approximation are:

- The main issues with the models tried are: (i) considerable overestimation or underestimation of emissions at both individual and aggregated level; (ii)

inconsistent values for some methods (i.e. negative emissions for the linear regression model); (iii) interpretability issues with the neural network model that complicate the explicability of the results.

- From the estimations obtained there is not a clear relation between the selected variables and the emissions. Consequently, the model obtained is not sufficiently adequate to be used for any type of estimations that could result in real economic consequences.
- Further refinements and fine-tuning could be addressed in order to improve the model by trying other alternatives/models, adding new variables or including new data for training.
- The comparison of the estimated values with real data results in considerable differences.
- The main variable that survives in the linear regression by steps is the Gross Value Added (GVA) of the company.

Within the various European initiatives aimed at developing climate-related indicators (including sustainable finance, physical risks and carbon emissions)⁷, a dedicated working group has recently been established focussed on estimating companies' carbon footprints. The Bank of Spain is actively participating in this subgroup with the objective of sharing its expertise and refining the existing model presented in this section in order to align with the approach agreed by the group.

4.2 Voluntary sustainability-related indicators in the annual accounts deposit in the Mercantile Registers

Since 1991 the Banco de España's Central Balance Sheet Data Office, the Mercantile Registers and the Accounting Regulatory Authority (Spanish Standard Setter, called ICAC) have been collaborating under a cooperation agreement to facilitate the legal deposit of the annual accounts using digital technology (XBRL files) in order to also promote the statistical use of the annual accounts that undertakings area legally required to file with the mercantile registry of the province in which their registered office is located.

This working and collaborative partnership between these institutions has enabled the exploration of new areas of information such as sustainability-related topics. Therefore, during the 2023 accounts filing campaign, a set of voluntary ESG indicators (based on indicators contained on the European Sustainability Reporting Standards) was introduced by Ministry of Justice, in a voluntarily framework, to improve the availability and reporting of standardized information on sustainability data by Spanish non-financial companies, with special emphasis on SMEs for which there is almost no information available of this kind (table 4 shows the specific list of ESG indicators selected).

⁷ For further information on climate change-related indicators please refer to the following link

Environmental indicators			
Sustainability indicators (voluntary content)			
	Unit of measurement	Value	Amount (€)
Scope 1 emissions	tonnes of CO ₂ eq		
Scope 2 emissions	tonnes of CO ₂ eq		
Scope 3 emissions	tonnes of CO ₂ eq		
Energy consumption within the organization	MWh		
Water consumption	m ³		
Social indicators			
		2022	2021
Percentage of women in the board	%		

Table 5. Voluntary sustainability-related indicators included in the annual accounts deposit in the Mercantile Registers

This information, when possible, has also being included in monetary terms to make it easier (especially for SMEs) to report it as they are more used to reporting accounting items in monetary terms and normally, is more feasible to obtain it directly from their accounting areas.

This initiative was conceived as a way to lead the path towards homogenized and digital reporting in order to increase corporate transparency, and contribute to the disposal of more granular information (with company-by-company data and not just at consolidated level which is mainly the type of data currently available).

During the 2023 campaign (with data of 2022) both the quantity and quality of information reported has not been great (specially with environmental indicators), however, it's just the first year this information was asked for and on a voluntary basis. Additionally, a possible explanation for these initial results could be that when this information was requested (in 2023), the CSRD was not yet in effect, and therefore, the calculations of these indicators by companies were not mature. Presumably in future campaigns results will be more encouraging and positive.

4.3 Estimation methodology of energy consumption and GHG emissions based on specific items from the P&L account

Another conceivable approach explored by the Banco de España's Central Balance Sheet Data Office to bridge SME ESG data disposal data gap is the use of specific accounting items from the P&L account to try to estimate environmental indicators. The information received on a voluntary basis from the undertakings in the Central Balance Sheet Data Office includes complementary data and, if necessary, the annual accounts for further detail consultation.

The request for this information, through the questionnaire by which economic and financial information is collected from medium and large companies, has served as a pilot to test to what extent these types of companies were already preparing to supply this information during a period when the corresponding regulation (CSRD) was being developed. This information is being asked since 2021 and approximately is available for 200-300 companies per year (approximately 30% of companies asked provide this information).

Under this situation, two specific items were gathered in order to measure the environmental impact of the undertaking. The first one is related to trying to estimate the impact in terms of GHG emissions by using the fuel expenses included in the operating expenditure (OPEX) and the fuel consumption included in the cost of goods. With these two items it is possible to approximate a potential environmental impact of a company and compare it among sectors in terms of these expenses. On the other hand, the second item is related with the electricity included in the operating expenditure (OPEX) which could be used as a possible measure of the potential energy consumption in combination with fuel consumptions of the company itself.

A second derivative of the use of this information, is the possibility of combining the data collected in monetary units with information of average prices of electricity or diesel during a year in order to make an approximation of energy consumption in MWh and GHG emissions in TnCO₂eq (this could be seen as a possible alternative to have data on carbon footprint scope 1-2 at company level when data is missing). Unfortunately, there are some caveats when doing this approximation in terms of what type of source of electricity is being used, the type of fuel consumed... that cannot be incorporated in the calculus and, consequently, this results in a less relevant approach for direct estimation. Meanwhile, it could be viewed as a good measure and source of contrast for evaluating environmental performance of companies and, especially, when there is no other information available as is the case for the majority of SMEs.

Year 2021	Fuel consumption (000€) from P&L		Scope 1 emissions (000 TnCO ₂ e) from sustainability reports	
	Aggregated value	Average value per company	Aggregated value	Average value per company
ENERGY	532.131	25.340	43.778	3.989
INDUSTRY	184.467	3.925	22.865	1.046
RETAIL & FOOD SERVICE	28.709	1.025	784	107
INFORMATION & COMMUNICATION	3.854	642	200	25
CONSTRUCTION & REAL ESTATE	1.184	118	869	114
TRANSPORTATION & STORAGE	549.871	32.345	11.961	2.071
REST	1.538	40	90	33

Table 6. Example of comparison between P&L fuel consumption and Scope 1 emissions for a sample of non-financial corporations in year 2021

This accounting information may be used as a source of contrast for information reported by companies to see if there is a coherent evolution between ESG indicators reported in the sustainability statements (i.e. scope 1 emissions or energy consumption) and the information in monetary terms included in the Profit and Loss account (i.e. electricity or fuel expenses). Although their behaviour does not need to be directly homogeneous, it is expected in most cases to have a similar evolution pattern since it is understood that, for example, higher electricity expenses should result in higher energy consumption reported by the company or higher fuel consumption expenses should result in higher scope 1 emissions. For example, in table 6 there is an example that presents a similar behaviour between sectors when comparing fuel consumption from the P&L of companies in euros with scope 1 emissions in TnCO₂e from the sustainability statements of companies, in both cases sectors such as energy or transportation are the ones that present higher values both in aggregated and relative terms.

It is possible that other exogenous variables and circumstances could explain a completely different behaviour between accounting items and environmental indicators but, normally, the behaviour and evolution should have some correlation and the accounting information helps considerably to this type of analysis.

In summary, obtaining climate-related data from SMEs poses challenges due to the absence of direct regulatory mandates and the limited familiarity of SMEs with such data. To address this, alternative approaches leveraging on information that SMEs are familiar with—such as energy or fuel consumption bills—are essential for generating climate-related indicators. This becomes especially crucial as banks increasingly demand and compel SMEs to provide this information. Currently, banks assist SMEs in estimating climate-related indicators using accounting items like energy or fuel consumption bills.

4.4 Data-sharing: Public-private collaboration for the creation of a Green Data Space

Relevant agents in both public and private areas have access to specific climate-related information of companies (including SMEs). The value added of this information on its own is limited but, when combined, the utility and usefulness increases exponentially. The interoperability of this information enables to better anticipate and simulate the economic impacts of climate change and prevent its negative effects.

For this reason, under the umbrella of the BIDA observatory⁸, a public-private data collaboration project arise in order to try to create a Green Data Space for data sharing among different public and private institutions. The objective is to create a secure public-private Green Data Space where data can be combined and interoperate for the generation of value added and support the definition of economic policies as well as for any other interesting use cases, preserving confidentiality of the data.

The structure to develop this Green Data Space could be the following (all these lines are currently under evaluation by the BIDA members):

- **Data:** relevant information on the field provided by partners (i.e. companies energy and water consumption based on the bank bills provided by financial institutions, satellite images and geolocalization information provided and mobile position and mobility data by a telecommunications companies, economic impact occurred due to climate change events obtained by insurance companies...).
- **Platform & Infrastructure for confidential usage:** safe and secure platform in the cloud or on-premise, using PET technologies or federated data, and aligned with the European data strategy with a clear governance procedure.

⁸ The Spanish observatory BiDA on Big Data, Artificial Intelligence and Data Analytics is a forum founded in 2018 designed to address the intersection between business and artificial intelligence, providing valuable information on the changing landscape of this technology and the impact on various industries and companies in Spain. The forum is composed by organisms both from the public and private area.

- **Analytics & AI:** partners provide large experience in AI projects focussing on data interoperability, ML algorithms and ethical use of data, PET technologies and federated data solutions.
- **Insights:** preparation of interactive and visual dashboards to help decision making.

This ongoing initiative would aim to establish a secure and controlled environment for valuable climate-related information. Partners can maximize its use by leveraging on specific use cases and employing advanced analytical techniques to exploit this data. The ultimate objective is to assist both private and public companies in better anticipating and modelling the economic effects of climate change on households and firms. This initiative focuses on real-time monitoring of policy impacts related to climate change. One of the primary goals of the Data Gaps Initiative⁹ Recommendation 13 on access to private sources of data, administrative data and data sharing is to promote and facilitate such initiatives through common frameworks and clear agreements. These agreements, once defined, would allow public and private companies to share their data and expertise, leading to better decision-making.

This project would also be particularly relevant in order to have access to climate information about SMEs that is not publicly available and that large corporations related to the banking or telecommunications industry manage internally. With this information it is easier to correctly understand the evolution of the impacts and the situation of SMEs regarding climate change. Additionally, by sharing not only data but also estimation and imputation methodologies, the synergies in terms of increasing data quality and quantity are notable.

Reputation purposes, win-to-win use cases, knowledge sharing and data contrast or complementation are the main key drivers which guarantee the success of this type of public-private agreements.

4.5 Data sharing: the relevance of establishing an integrated repository of pollutants and GHG emissions among key Spanish public institutions

Another alternative approach (currently under evaluation) to fill SME data gaps is the possibility of building an integrated repository of pollutants and GHG emissions among the main public registers of a country. In this case, the example is focused on Spain but the idea could be extrapolated to any jurisdiction that could enable it.

In Spain there are different regulatory bodies that collect and receive environmental information, since 10 years ago, for different reasons, mainly driven by industrial data. Taking into account the possible synergies between registers and the exponential relevance this type of information has been gaining in recent years, a project for building a common national repository of GHG data was projected although has not yet started among these institutions.

⁹ In the fall of 2021, the G20 FMCBGs endorsed the concept note for the third phase of the initiative (DGI-3). DGI-3 includes 14 recommendations addressing four statistical areas (i) climate change; (ii) household distributional information; (iii) Fintech and financial inclusion; and (iv) access to private sources of data and administrative data, and data sharing [link]

The idea is to incorporate under a common framework and structure all the different national GHG and pollutant registers for better data usage and comparability with the information provided by companies in their sustainability reports (CSRD from 2025 onwards). Additionally, it aims to achieve the best alignment as possible with the future regulatory requirements for ESG data from companies in Europe such as the EU Corporate Sustainability Reporting Directive (i.e. if the repository contains Scope 1 emissions from companies, this approach could try to ensure that the metric included and reported is tons of CO2 equivalent as stated in the European regulatory requirements).

This future project (if materialised) will consist of a joint collaboration among different administrative registers and institutions; Bank of Spain, who will act as project promoter, data provider and sustainability expert adviser, and the National Statistics Institute (NSI or INE in Spanish acronym) who will act as project promoter and leader and also collaborates with national registers which are data providers. The project's goal would be to create an integrated repository to store different climate-related micro databases available in each of the register for statistical purposes, comparing this register data with the climate information provided by the companies in its ESG reports. This facilitates the use and the interoperability of the data increasing value added for stakeholders.

The project was organised in two phases. The first phase developed in 2022 and 2023, consisted in providing feedback and organizing meetings by the different participating institutions in order to better understand the functioning, characteristics and methodologies behind these databases. The second phase (currently under discussion) will consist on the preparation of an integrated repository with the metadata associated¹⁰. For the moment, a preliminary taxonomy with the variables that should contain this common repository have been agreed (table 7 reflects the variables selected for this purpose).

PRELIMINARY TAXONOMY DEFINED
National Tax Identifier
Internal code of the company within the administrative register
Company name
Year
Variable/indicator (i.e. Scope 1, Energy consumption...)
Value
Unit of measurement (i.e. TnCO2e, MWh...)
Type of indicators (i.e. GHG, Energy, Water)
Data source (i.e. sustainability statements, EU ETS, PRTR...)
Administrative Register that provides the data
Frequency (i.e. annual, monthly...)
Granularity (Industrial complex / company level / Consolidated level...)
Name of the industrial complex (when available)
Geolocalization of the industrial complex (when available)
Confidentiality (yes/no)
Level of use (statistics purpose, publicly available...)

¹⁰ Level of granularity, frequency of update, unit of measurement, degree of confidentiality, quality degree...

PRELIMINARY TAXONOMY DEFINED
Quality level (high, médium, low...)
Original or transformed data
Commentaries

Table 7. Preliminary taxonomy agreed for the common repository

These different databases, currently available, are managed by Bank of Spain, National Statistics Institute (INE) and the Ministry for the Ecological Transition and the Demographic Challenge-Energy (MITERD), in particular the following databases have been under study:

- PRTR-E (Spanish Pollutant Release and Transfer Register) (*"Registro Estatal de Emisiones y Fuentes Contaminantes"*) which is responsible for collecting information (at industrial complex level) on emissions to the atmosphere, water and soil of polluting substances and data on waste transfers from the main industries.
- EU ETS (Emissions Trading System) – National Registry of GHG rights (RENADE) (*"Registro Nacional de Derechos de Emisión de Gases de Efecto Invernadero"*) where information related to direct GHG emissions ("verified emissions") is reported as well as information on GHG emissions rights trading.
- Carbon Footprint Register (*"Registro de huella de carbono, compensación y proyectos de absorción de dióxido de carbono"*) which is a public register under MITERD supervision that collects information related to the carbon footprint of Spanish firms that report on a voluntary basis.
- Environmental accounts from National Statistics Institute (*"cuentas medioambientales del INE"*) present data regarding the polluting emissions into the atmosphere, in a way that is compatible with the National Accounts System, registering the emitting agents, broken down by branch of economic activity and households as final consumers.
- Spanish Inventory System (*"Sistema Español de Inventario"*). The MITERD, using IPCC (Intergovernmental Panel on Climate Change) and EMEP/EEA methodology with the NFR/CRF (Nomenclature for Reporting / Common Reporting Framework) nomenclature, prepares the national emissions inventory that groups emissions into sectors, categories and subsectors.
- Sustainability reports from companies. Companies that meet the requirements of Law 11/2018 (which will gradually be substituted by the CSRD) are obliged to prepare a sustainability report where they must report ESG information, including the carbon footprint measured through Scope 1, 2 and 3 emissions.

All these databases shared the common feature of containing information related to Greenhouse Gases produced under different regulations or standards. That causes the coexistence of different databases that are difficult to reconcile. The main problems for comparing and creating a common and interoperable national directory with these databases arise due to:

- The different level of granularity of each of them (data is available at industrial complex, aggregated level or company/group level),
- The degree of confidentiality of the information (i.e. the company id identifier) which complicate the matching process and constrains data

sharing between institutions due to jurisdictional and internal policies limitations.

- The different metrics (Scope 1, Scope 2, Scope 3, total GHG emissions...) and units of measurement (CO₂e, by gases...) used of each database.
- The different thresholds (i.e. information not reported if total installed capacity of industrial complex is less than 25MW...) and the limited number of activities each of them contains (i.e. just industrial and air activities...).

However, although there is a comprehensive list of difficulties that need to be taken into consideration when mixing and making this information interoperable, the benefits would be considerable in terms of:

- The possibility of comparison and contrast between sources.
- Quality control checks for increasing robustness of each database.
- Complement information between different sources for a better and more reliable image of the GHG and pollutants of the companies.
- Statistical access to a common and single integrated repository for a better and more agile exploitation of this micro data level for users.

Consequently, the main lesson to extrapolate from this experience is the importance of creating a common framework where administrative registers can collaborate and share data in order to maximize utility for users and society.

5. Conclusions

Small and Medium enterprises (SMEs) are currently facing a controversial moment as they are starting to suffer the consequences of not preparing ESG information and addressing climate-related risks. From a regulatory perspective they are mostly exempted from any type of direct requirement but, indirect drivers such as the value chain requirements from bigger companies, the access to financing products and the new requirements from markets and investors are forcing them to prepare and address these new risks in order to survive in this new competitive environment where ESG risks, policies, impacts and actions are being considered. These process will not be easy for SMEs as the majority are not prepared and lack technical capacity and resources to elaborate and analyse this type of information.

The Central Balance Sheet Data Office of the Banco de España, aware of the increasing importance of granular ESG data from SMEs, is continuously seeking alternatives and proxies to minimize this data gaps as the ones presented in this paper (public-private data sharing agreements, econometric models, voluntary indicators, national public GHG repository or P&L accounting items). This will help to increase the quality and quantity of ESG data and will be beneficial not only for regulators, bigger companies, society and the financial sector but also for SMEs that are willing to evolve and adapt to these new circumstance and challenge which results in numerous risks but also opportunities in the future.

References

- [1] SEBEN OZKAN, SILVIA ROMAGNOLI AND PIETRO ROSSI. "A novel approach A novel approach to rating SMEs' environmental performance: Bridging the ESG gap" (2023)
- [2] HEBAH SHALHOOB AND KHALED HUSSAINEY. "Environmental, Social and Governance (ESG) Disclosure and the Small and Medium Enterprises (SMEs) Sustainability Performance" (2022)
- [3] AMIR GHOLAMI, PETER A. MURRAY AND JOHN SANDS. "Environmental, Social, Governance & Financial Performance Disclosure for Large Firms: Is This Different for SME Firms?" (2022)
- [4] PAULA GOMEZ DOS SANTOS, FÁBIO ALBURQUEQUE, MARIA ALBERTINA BARREIRO RODRIGUES AND ANA ISABEL MORAIS. "The Views of Stakeholders on Mandatory or Voluntary Use of a Simplified Standard on Non-Financial Information for SMEs in the European Union" (2022)
- [5] ANDREA GIACOMELLI. "EU Taxonomy for non-financial undertakings: summary reporting criteria and extension to SMEs" (2021)
- [6] BORJA FERNÁNDEZ-ROSILLO SAN ISIDRO, EUGENIA KOBLENTS LAPTEVA Y ALEJANDRO MORALES FERNÁNDEZ: Micro-database for sustainability (ESG) indicators developed at the Banco de España (2022).
- [7] NGFS EN on DATA: "Information note on Building National Data Directories" (2023)
- [8] NGFS EN on DATA: "Information note on Improving Emissions Data" (2023)
- [9] EFRAG. "Draft EFRAG Voluntary ESRS for non-listed Small and Medium-Sized Enterprises – Exposure Draft" (2023)
- [10] EUROPEAN PARLIAMENT AND COUNCIL. "Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting" (2022)
- [11] GLOBAL REPORTING INITIATIVE (GRI) documents published on their website related to: "Small business big impact: SME sustainability reporting from vision to action" and "Empowering Small Business"

ADDRESSING SUSTAINABILITY DATA GAPS FOR SMES: CHALLENGES, ALTERNATIVE SOLUTIONS AND DATA- SHARING

Borja Fernández-Rosillo San Isidro

Central Balance Sheet Data Office
Statistics Department

**12TH IFC CONFERENCE ON “STATISTICS AND BEYOND: NEW DATA FOR DECISION MAKING IN
CENTRAL BANKS”**

BASEL
23th August 2024

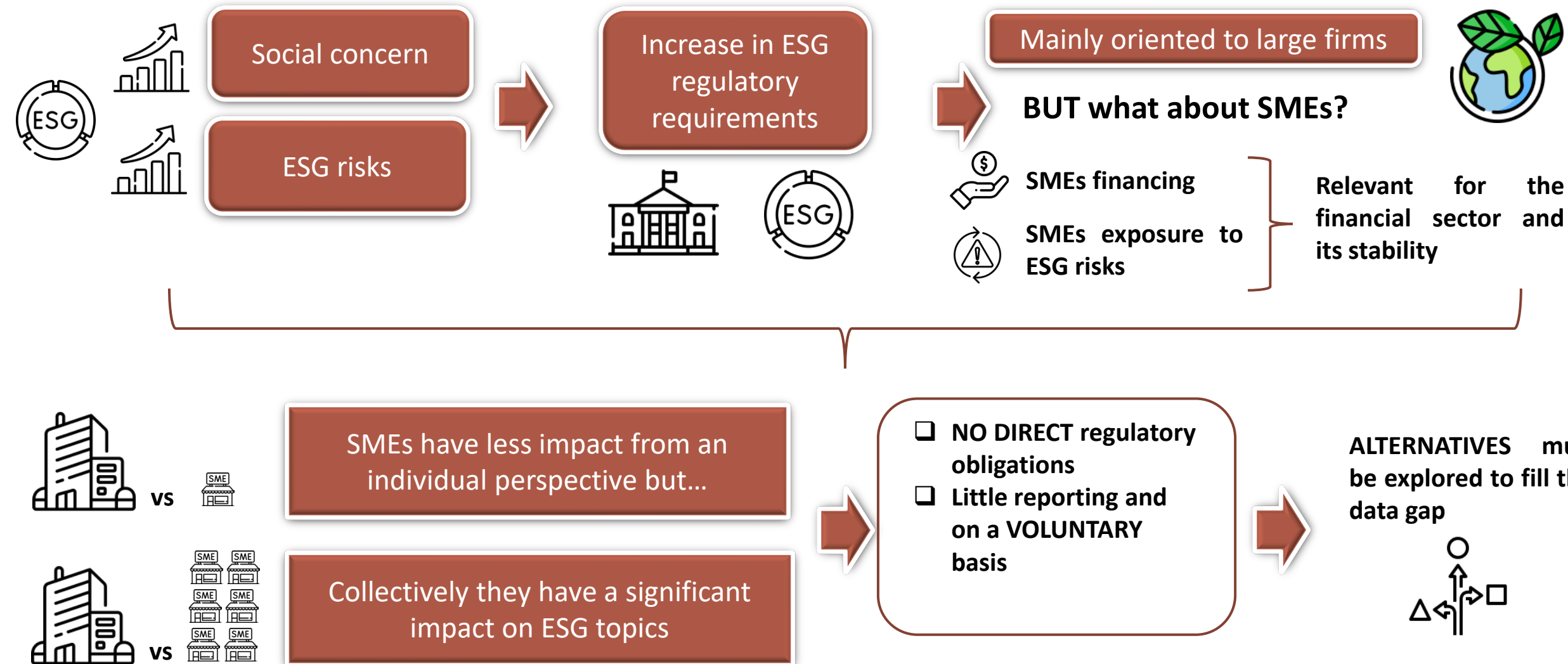


INDEX

1. Introduction and context
2. Current situation of ESG indicators for SMEs
3. Stakeholders needs and main drivers to address them
4. Alternative measures to fill the data gap on SMEs ESG:
 - 4.1 An econometric model for estimating the carbon footprint [work in progress]
 - 4.2 Voluntary sustainability-related indicators in the annual accounts
 - 4.3 P&L items as a proxy for energy and GHG emissions
 - 4.4. Data-sharing: Public-private collaboration for the creation of a Green Data Space [work in progress]
 - 4.5 Data-sharing: Establishing an integrated climate-data repository ¿?
5. Conclusions

1. INTRODUCTION AND CONTEXT

SME ESG data

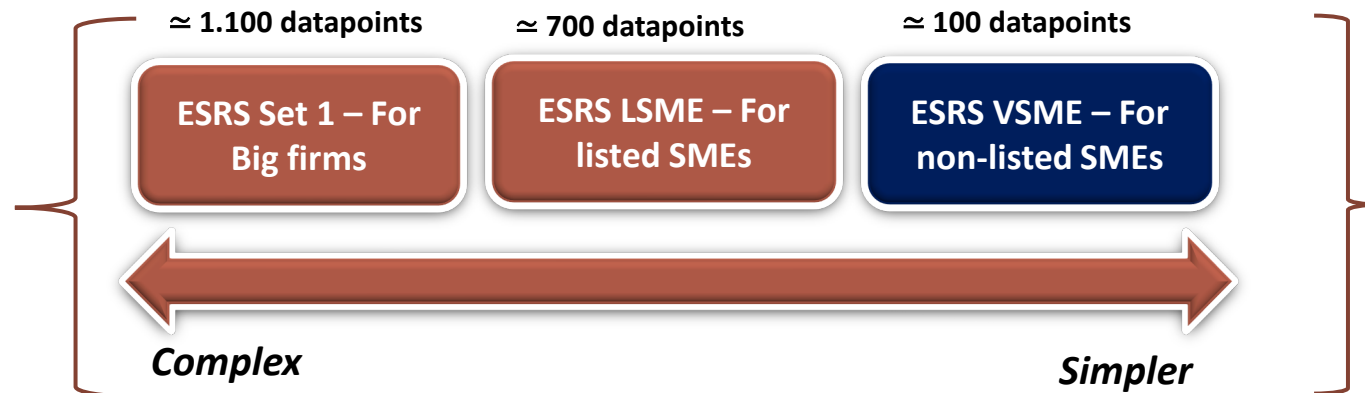


2. CURRENT SITUATION OF ESG INDICATORS FOR SMES

Landscape of ESG requirements for SMEs



- ❑ Only EFRAG has developed SME oriented standards but are VOLUNTARY



- ❑ NO obligatory requirements



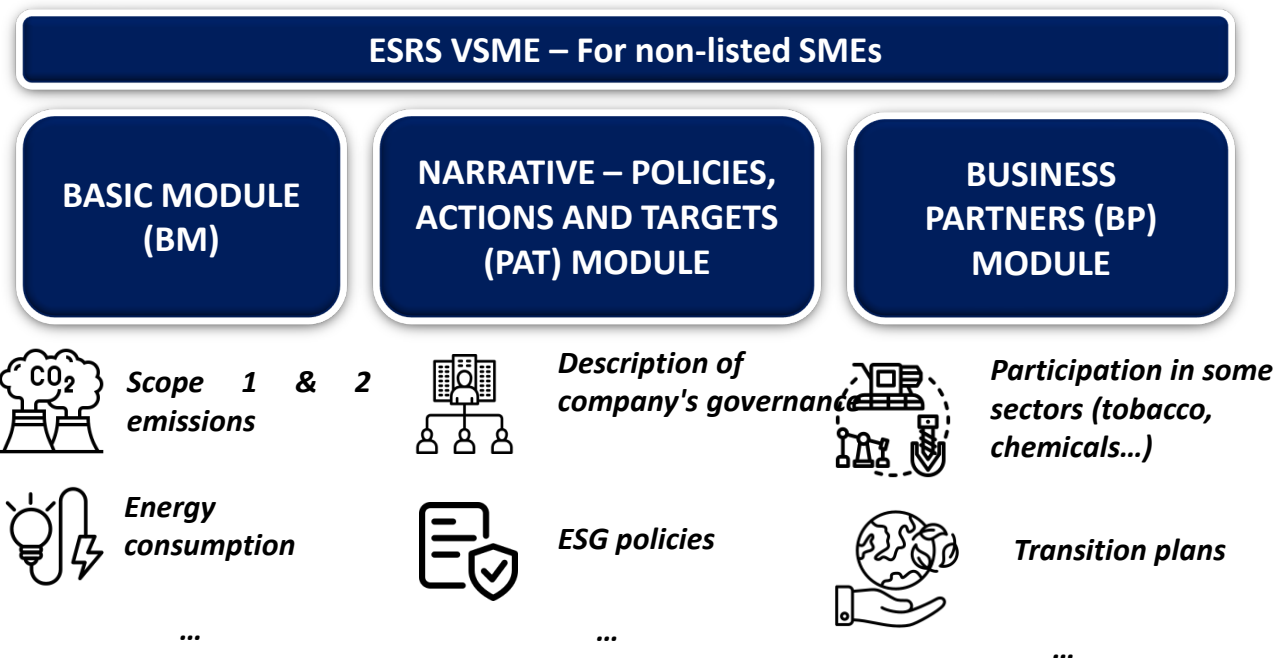
- ❑ Few ESG SMEs oriented standards developed



- ❑ Most SME lack resources and technical capacities to prepare this information



Examples



3. STAKEHOLDERS NEEDS AND MAIN DRIVERS TO ADDRESS THEM

Indirect drivers for SME increasing ESG data

Non-listed SMEs are not required to report sustainability information under the CSRD but...

INDIRECTLY they will be “pushed” to present this information due to:

Access to financial services

+99% of total firms in Spain are SMEs

Value chain requirements



Market transparency

For this reason, voluntary indicators for non-listed SMEs have been developed by EFRAG (VSME)



4. ALTERNATIVE MEASURES TO FILL THE DATA GAP ON SMES ESG

4.1 An econometric model for estimating the carbon footprint [work in progress]

Objective

To build an econometric model capable of estimating the carbon footprint of each Spanish non-financial company based on the greenhouse gas (GHG) emission intensity of its sector of activity (based on the NSI environmental accounts) in combination with other accounting variables

Response variable (yt)

Emission coefficient (TnCO₂eq per 1,000€ produced)

Explanatory variables (xt)

Average age of fixed assets in years

NACE at 2 digits

Years

Size proportion

Average number of employees

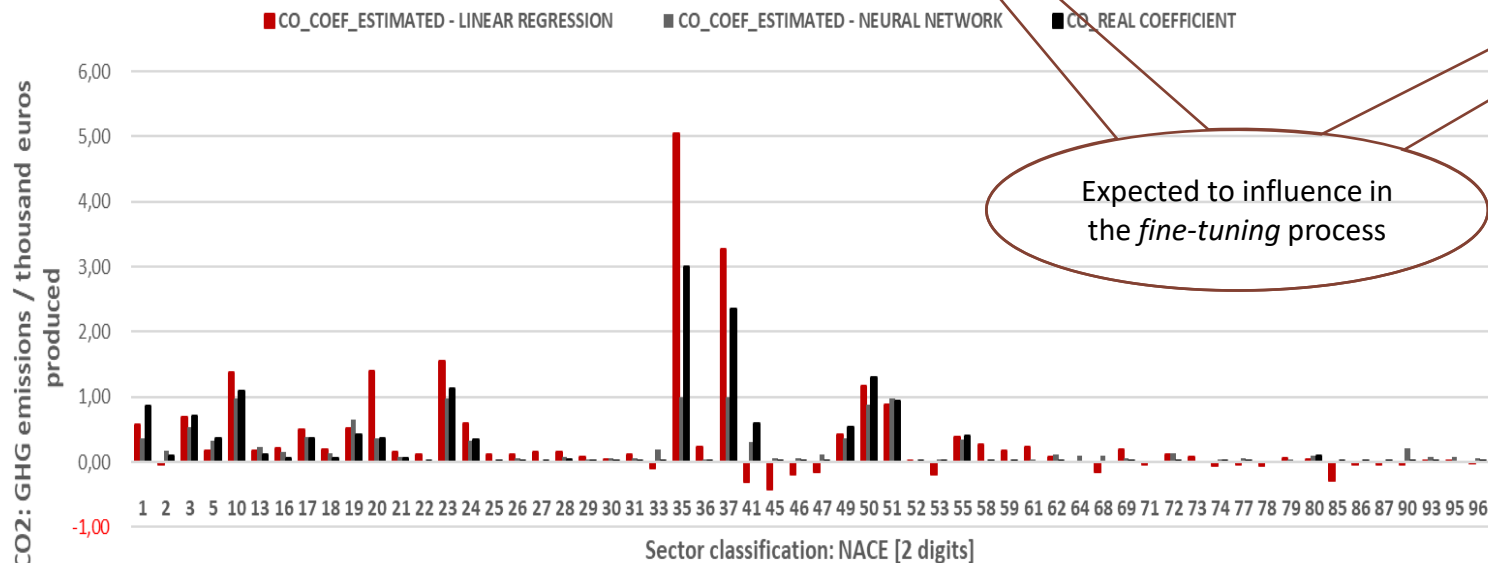
Accounting variables in relative terms

- Net amount of turnover
- Depreciations
- Sales
- Intermediate consumptions

- Tangible assets
- Intangible assets
- OPEX
- Personnel costs
- Gross Value Added
- Current assets
- Suppliers



COMPARISON OF ESTIMATION METHODS FOR SECTOR AGGREGATED DATA: LINEAL REGRESSION VS NEURAL LINK VS REAL DATA



Different models tried (linear regressions, neural network...)

GVA and NACE sector as the most significant explanatory variables

CONCLUSIONS:

- ☐ Considerable overestimation or underestimation of emissions both at individual and aggregated level
- ☐ Inconsistent values for some methods (i.e. negative emissions for the linear regression model)
- ☐ Interpretability issues with the neural network model that complicate the explicability of the results



4. ALTERNATIVE MEASURES TO FILL THE DATA GAP ON SMES ESG

4.2 Voluntary sustainability-related indicators in the annual accounts

Set of **VOLUNTARY**
ESG indicators on the
annual accounts
deposit



On a **STANDARDISED**
format and aligned
with the CSRD



- ☐ Increase quantity and quality of data
- ☐ Help firms (specially SMEs) start familiarizing with this reporting



- ☐ First results for environmental indicators are **extremely poor both in quantity and quality terms**



- ☐ Although it is the first year, **voluntary information request seems to not work**



- ☐ Consequently, pressure from both banks and the value chain is essential to obtain this data



Environmental indicators



Sustainability indicators (voluntary content)

	Unit of measurement	Value	Amount (€)
Scope 1 emissions	tonnes of CO2eq		
Scope 2 emissions	tonnes of CO2eq		
Scope 3 emissions	tonnes of CO2eq		
Energy consumption within the organization	MWh		
Water consumption	m3		

Social indicators

		2022	2021
Percentage of women in the board	%		

4. ALTERNATIVE MEASURES TO FILL THE DATA GAP ON SMES ESG

4.3 P&L items as a proxy for energy and GHG emissions



Most SMEs are not familiar with climate-related indicators...



BUT they surely are with accounting items in monetary terms



Accounting items from the P&L account

- ☐ OPEX – Fuel expenses (€)
- ☐ OPEX – Electricity expenses (€)
- ☐ Cost of Goods – Fuel consumption (€)

1

- ☐ Replicate climate-related indicators behaviour
- ☐ Source contrast

of

2

- ☐ Estimate GHG emissions or energy consumption (in MWh) based on average prices

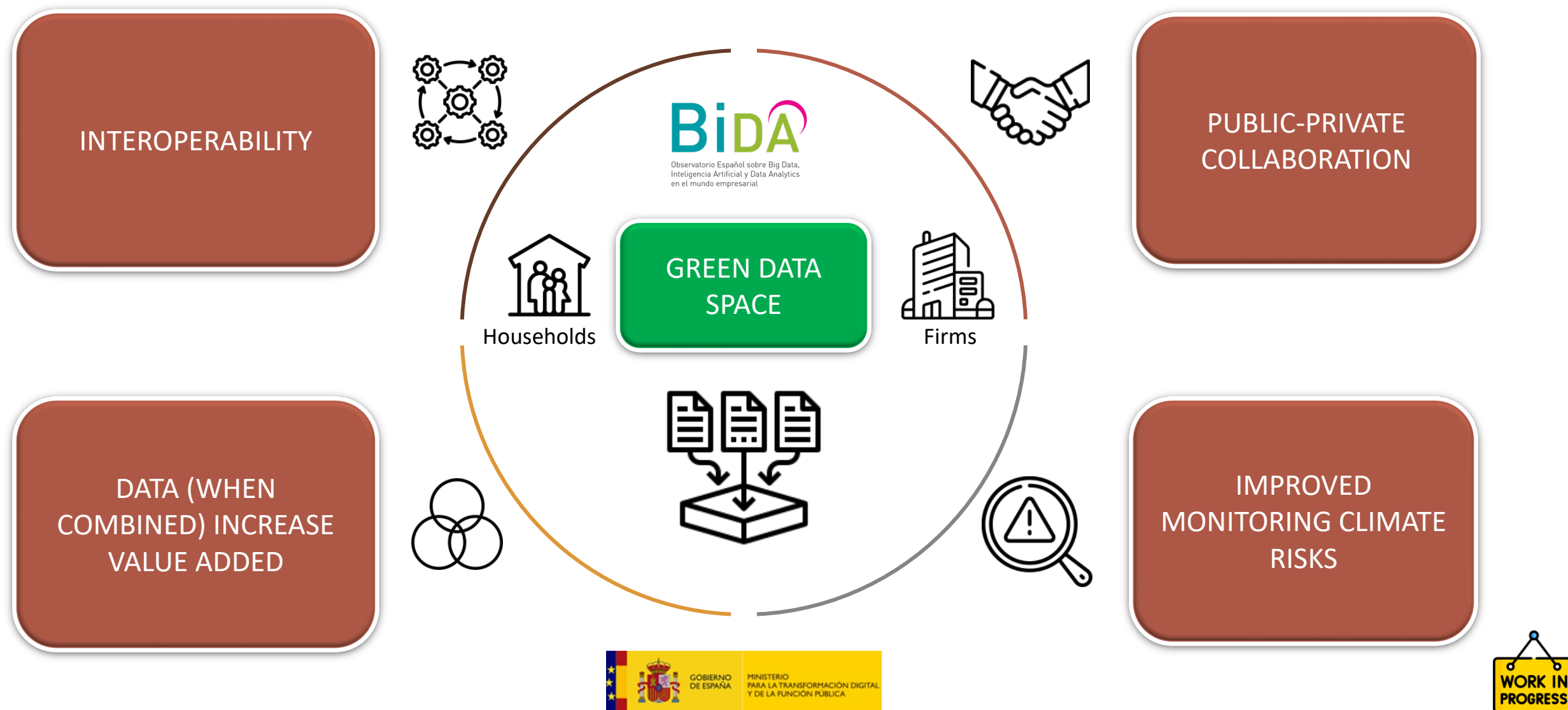
Example of comparison between fuel consumption (000€) from the P&L accounts with Scope 1 emissions (000 TnCO₂e) from sustainability reports

Year 2021	Fuel consumption (000€) from P&L		Scope 1 emissions (000 TnCO ₂ e) from sustainability reports	
	Aggregated value	Average value per company	Aggregated value	Average value per company
ENERGY	532.131	25.340	43.778	3.989
INDUSTRY	184.467	3.925	22.865	1.046
RETAIL & FOOD SERVICE	28.709	1.025	784	107
INFORMATION & COMMUNICATION	3.854	642	200	25
CONSTRUCTION & REAL ESTATE	1.184	118	869	114
TRANSPORTATION & STORAGE	549.871	32.345	11.961	2.071
REST	1.538	40	90	33

- ☐ However some caveats must be addressed in terms of type of source of electricity, type of fuel, energy mix... for a fine-tune approximation

4. ALTERNATIVE MEASURES TO FILL THE DATA GAP ON SMES ESG

4.4 Data-sharing: Public-private collaboration for the creation of a Green Data Space [work in progress]



4. ALTERNATIVE MEASURES TO FILL THE DATA GAP ON SMES ESG

4.5 Data-sharing: Establishing an integrated climate-data repository ¿?



Sistema Español de Inventario (SEI)

Spanish Inventory System

CLIMATE ACTION
European Union Transaction Log

EU ETS (Emissions Trading System)
National Registry of GHG rights

PRTR
España

Registro Estatal de Emisiones y Fuentes Contaminantes

PRTR-E (Spanish Pollutant Release and Transfer Register)

Carbon Footprint Register

INē
Instituto Nacional de Estadística

Environmental accounts

BANCODE ESPAÑA
Eurosistema

Sustainability Reports from companies

Effort to share all the existing info from GHG in different Spanish institutions



Milestones achieved



- ☐ Technical Note
- ☐ Preliminary taxonomy for the integrated repository

Future work*



- ☐ Solve confidentiality issues
- ☐ Build the architecture for the repository

PRELIMINARY TAXONOMY DEFINED
National Tax Identifier
Internal code of the company within the administrative register
Company name
Year
Variable/indicator (i.e. Scope 1, Energy consumption...)
Value
Unit of measurement (i.e. TnCO2e, MWh...)
Type of indicators (i.e. GHG, Energy, Water)
Data source (i.e. sustainability statements, EU ETS, PRTR...)
Administrative Register that provides the data
Frequency (i.e. annual, monthly...)
Granularity (Industrial complex / company level / Consolidated level...)
Name of the industrial complex (when available)
Geolocalization of the industrial complex (when available)
Confidentiality (yes/no)
Level of use (statistics purpose, publicly available...)
Quality level (high, médium, low...)
Original or transformed data
Commentaries

(*) subject to further discussion and agreements among the institutions collaborating

5. CONCLUSIONS

Main highlights

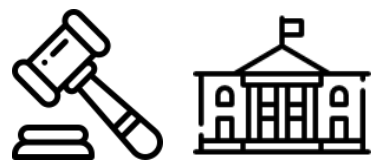
Sustainability-related information has become a reality

1

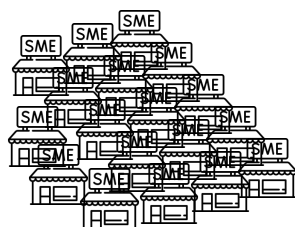


Efforts from regulatory agencies are focused on large companies BUT, what about SMEs?

2



CSRD



Although indirect drivers will force them to prepare this information...

3



Financial services

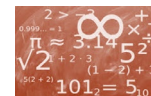


Value chain



Something needs to be done in the meantime

4



Econometric modelling



P&L accounting items



Voluntary reporting



Data-sharing agreements

In order to increase quantity/quality of ESG data for SMEs and necessary for all stakeholders involved

5



Increase data for better decision making



THANKS FOR YOUR ATTENTION

