IFC Workshop on "Addressing climate change data needs: the global debate and central banks' contribution"

6-7 May 2024

Measuring foreign direct investment carbon footprint: an experiment with micro data[1]

V Genre, A Magniez, D Nefzi and F Robin,
Banque of France

[1] This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the Central Bank of the Republic of Türkiye, the BIS, the IFC or the other central banks and institutions represented at the event.

Measuring FDI carbon footprint: an experiment with micro data

Véronique Genre, Alice Magniez, David Nefzi and François Robin[1]

Abstract

In response to the G20's Data Gaps Initiative (DGI-3), this paper develops a novel methodology for measuring the carbon footprint of Foreign Direct Investment (FDI) using firm-level data. While current approaches rely on macroeconomic data and input-output models, this study conducts a micro-level analysis, using granular emissions data from about 2,500 companies over the period 2016–2022. By focusing on FDI stocks rather than flows, the methodology offers an alternative assessment of greenhouse gas (GHG) emissions linked to both inward and outward FDI in France. Emissions data from listed companies is extrapolated to their foreign affiliates thanks to an allocation model based on firms' assets and sectoral environmental efficiency.

The findings reveal significant variations in the carbon footprint of French multinationals abroad and foreign-owned enterprises in France, with results strongly influenced by the environmental efficiency of different sectors and countries. The analysis also highlights substantial data gaps, particularly for unlisted firms and FDI in less-developed regions. The use of estimation models and proxies to address missing data illustrates both the potential and the limitations of micro-data approaches. Despite these challenges, the methodology provides a promising complement to existing macro-based methodology and opens new avenues for robustness checks. Future improvements in international data transparency, as well as regulatory initiatives like the European Corporate Sustainability Reporting Directive (CSRD), are expected to enhance the reliability and coverage of this approach.

Keywords: foreign assets, foreign direct investment, carbon emissions, environment, economic impact of globalization

JEL classification: C18, F21, F23, F64, Q51

[1] The views expressed here are those of the authors and not necessarily that of Banque de France. The authors thank Rocco Incardona and Jorge Diz Dias (ECB); Nadia Accoto, Giacomo Oddo and Daniele Di Pietro (Bank of Italy) for their helpful discussions at the "climate" FDI Network working sub-group; and Timothée Gigout (Banque de France) for paving the way in estimating the French FDI carbon footprint.

1. Introduction

In the aftermath of the Global Financial Crisis, back in 2009, the G20 Finance Ministers and Central Banks Governors first launched the Data Gaps Initiative (DGI) to fill in policy-relevant data gaps, which had prevented the crisis from being foreseen. The successful conclusion of the first initiative led to the launch of two subsequent phases with the objective to ensure methodological development, collection and dissemination of reliable and timely statistics for policymakers. In autumn 2021, the G20 endorsed a new guideline for the third phase of the initiative. DGI-3 included, among others, recommendations in the area of climate change and in particular, on how to measure FDI carbon footprint.

Countries generally welcome and even go to great lengths to attract inward Foreign Direct Investment (FDI). FDI is defined as an investment involving a long-term relationship, which reflects a lasting interest, of a resident entity in one economy (the so called DI, direct investor) in an entity that is resident in another economy (IMF, 2008). The lasting interest is determined when the direct investor acquires a minimum of 10% voting power or capital share of the investee (the direct investment enterprise, DIE). Compared with other forms of capital flows, FDI is less volatile. It has been associated with rising wages (Gopinath and Chen, 2003) and improved productivity, with technological know-how and skills' spillovers to the domestic economy (Sugiharti *et al*., 2022), with introduction of new industries and exports diversification (Tadesse and Shukralla, 2013), and, eventually and under the right conditions, with increasing domestic growth.

By contrast, the environmental impact of FDI remains unclear. Countries with low income tend to have lower pollution standards and may attract pollution-intensive FDI and see their greenhouse gases (GHG) emissions deteriorate. This is known as the *pollution haven* hypothesis, by which FDI significantly increases carbon emission levels in host economies. At the same time, FDI may deploy new technologies that are cleaner than domestic production and thus, foster improvements in the host economy. This is known as the *pollution halo* hypothesis. Empirical research has also found evidence of FDI accelerating the development and production process of a lower carbon economy thanks to FDI (see Gill *et al*. (2018) for a critical review of available empirical research to date).

One way to avoid having to take a stance on these two hypotheses would be to directly measure the carbon footprint of FDI. Yet this has proven challenging. Borga *et al*. (2022) suggest a methodology that may well set the standard for the coming years. This methodology is mostly based on input-output modelling, a method that has been gaining traction in recent years (Tukker *et al*. 2018, 2020). International input-output tables map all flows of goods and services across industries in different regions of the world from production to final demand. Together with GHG emissions by country and sector, one may solve linear systems to measure footprints of sectors, countries or a combination of both. Because of data availability, the IMF approach can only be applied to inward FDI flows for now.

In this paper, we experiment a method based on granular data using both inward and outward FDI stocks. Our method may be closer to FDI compilers' methods as it focuses on individual firms' data.

Our approach aims to complement the conceptual framework adopted by the IMF (Borga et al., 2022). While the IMF suggests a method to estimate carbon emissions of the share of total domestic production that is due to foreign-owned enterprises, their method relies on a particularly strong assumption. The carbon intensity of production among companies owned by non-residents across different sectors must be perfectly homogenous. In other words, it assumes no variation in production processes within industries. In contrast, our methodology leverages microeconomic data, particularly corporate greenhouse gas (GHG) emissions, to relax this uniformity assumption. This allows us to better capture the diversity in environmental efficiency across production processes. Furthermore, our approach focuses on invested companies involving a French resident, regardless of whether the investment is inward or outward.

Our methodology also diverges from the IMF's approach in another key dimension. The IMF primarily examines the activities of foreign affiliates of multinational enterprises, focusing on resident firms with at least 50% foreign ownership. By contrast, our scope is broader; encompassing all foreign direct investment (FDI) relationships, i.e. cases where a non-resident firm holds 10% or more of a resident firm's equity.

In the medium term, this granular approach will enable us to test the robustness of the assumptions underpinning the IMF's framework. Ultimately, we share a common objective: to evaluate the extent to which multinationals are responsible for GHG emissions through their FDI. This involves attributing accountability to the entity at the top of the group structure for GHG emissions across the entire ownership chain, including its subsidiaries and their sub-subsidiaries.

1. An approach based on granular data

To work out a FDI carbon footprint, our basic idea is to use GHG emissions corporate reports and to consider that companies that do not publish GHG reports but are from the same country, same sector of activity and have the same characteristics in terms of size, net revenues and fixed capital will broadly get the same environmental impact.

The Greenhouse Gas Protocol (GHG Protocol) is an international protocol providing a framework for measuring, accounting and managing greenhouse gas emissions resulting from operations of both private and public sector. In 2004, was released the Corporate Accounting and Reporting Standard that has now become the world's most widely used GHG accounting standard. It defines the concepts of scope 1, 2 and 3 GHG emissions. The International Sustainability Standards Board (ISSB) by the International Financial Reporting Standards (IFRS) Foundation requires companies to measure their greenhouse gas emissions in accordance with GHG Protocol[2].

Scope 1 includes direct GHG emissions from fixed or mobile sources controlled by the organization (e.g., direct combustion of fossil fuels to power heat sources or stationary combustion engines and emissions from all vehicles running on fossil

[2] https://ghgprotocol.org/blog/ghg-protocol-use-within-issbs-ifrs-s2-standard-enables-widespread-adoption-common-standard-ghg

fuels). Scope 2 covers GHG emissions arising from purchased energy — from a utility provider. That is, all GHG emissions released into the atmosphere, originating from the consumption of purchased electricity, steam, heat and cooling. Ultimately, Scope 3 includes indirect emissions not owned and not included in Scope 2. These emissions occur in the value chain of the company reporting the data and include both upstream and downstream emissions.

The GHG Protocol defines methods for setting boundaries for a GHG inventory: these determine which entities (e.g., subsidiaries, joint ventures, partnerships) and assets (e.g., facilities, vehicles) will be included in the scope 1 and scope 2 GHG emissions inventory. When an organization chooses an approach to consolidate GHG emissions, among three options, it has to apply it consistently to define entities and assets to include in their reporting[3]. This can yield some heterogeneity among international groups' GHG reports. Hereafter however, we assume that companies follow consolidation method through full integration accounting.

A first question arises as to which GHG emission perimeter to consider. Focusing on Scope 1 carbon footprint should prevent double counting, but it limits firms' accountability. Having in mind Scope 1 emissions for an energy-producing firm are the Scope 2 emissions of a firm in the manufacturing sector, Scopes 1 and 2 carbon footprint is the most common definition that encompasses GHG emissions resulting from firms' activities. Adding Scope 3 emission implies more over-counting (Charpentier *et al.*, 2023), but this is also a representation of what a firm is accountable for. We have experimented with all three scopes. Scope 3 results, however, are much less reliable. Indeed the corporate standard does not require Scope 3 emissions to be presented so that much fewer firms actually report these. As it requires collecting data at product-level from suppliers, many firms only report a few Scope 3 categories at their discretion, which leads to inaccurate estimations of their overall Scope 3 emissions.

In France, all companies with more than 500 employees have had to publish their GHG emissions on a yearly basis for the past 12 years, but only half of them actually comply. However, an increasing number of firms report GHG total emissions forced by either law, political pressure or financial incentives. Several private data providers, such as Institutional Shareholder Services (ISS), Carbon 4 Finance, the Carbon Disclosure Project, Refinitiv or Institutional Shareholder Services, now collect this data directly from companies' reports or estimate it. In total, Scopes 1 and 2 GHG emissions can be extracted for 28,000 companies worldwide, 20% of which being provided by the company itself and 80% being estimated by data providers, not always with a fully transparent methodology. This data always represents consolidated GHG emissions at the group level.

An iterative process, depending on data availability

Our aim is to measure the carbon footprint of outward and inward French FDI. To do so, we experiment with those DIEs being part of in an FDI framework involving a French resident, regardless of the investment direction.

[3] https://www.epa.gov/climateleadership/determine-organizational-boundaries

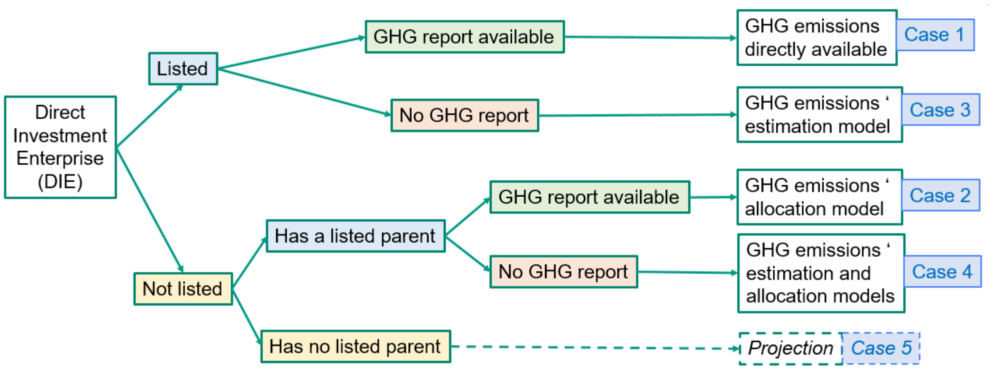
Our challenge here is to allocate the whole group's GHG emissions to its different affiliated entities (DIEs). Overall, our approach can be summarised in two steps:

1) First, we gather GHG total emissions for the group from ISS;

2) Then, we allocate this total amount to its different DIEs according to the models depicted below.

These two steps largely depend on data availability. A decision tree can summarise the whole process (Figure 1). When a company is listed, data is more widely available, be it detailed financial statements, size, financial ratios, sector of activity, etc. For those companies, we generally have GHG emissions data that can be used directly (Case 1). However, for the rare listed DIEs where GHG data is not available (Case 3), we need to estimate it. All links with the DIE are determined by the direct investor's shareholding.

Figure 1 – Decision tree for choosing the GHG emissions estimation method



One important issue to bear in mind is that DIEs are often unlisted: in the case of France, only 10% of FDI assets and 20% of liabilities (in value) are listed on a stock exchange. Because GHG data is only available for listed companies, we look for listed parents of unlisted DIEs along the ownership chain, be it the direct investor or the ultimate parent. To do so, we rely on the LIFI database, maintained by the French statistical office, INSEE. This database maps ownership links between companies domestically, but also abroad since it can identify foreign parents and group heads. Hence, we get a list of possible information-rich parents for each of our DIEs[4].

We then look for the ISIN code of these parent companies using the European Centralised Securities DataBase (CSDB), either directly, thanks to the SIREN code (French company's unique national identifier) or indirectly, thanks to text mining techniques running through companies' names. This latter process requires a manual checking step. Eventually, we can match a fair share of the list of unlisted DIEs with a relevant listed parent higher up in the chain of ownership. GHG total emissions' data is available for more than 90% of the listed parents and because the parent is listed,

[4] For inward FDI, DIEs are domestic and listed parent companies are generally located abroad. For outward FDI, DIEs are foreign and parent companies is usually domestic, but it may happen that the resident company is only a link in the whole ownership chain. Thankfully, the French economy does not shelter SPEs (Special Purpose Entities), which simplifies the view over the whole ownership chain.

we also have access to all kind of additional information that will helps us design a GHG allocation model for its DIEs (Case 2).

For those rare cases where no listed parent, - either direct investor or at the top of the chain of ownership -, reports any GHG emissions, we need an estimation model: first, we estimate GHG emissions at the group level for the listed parent identified with an ISIN code. Then, we allocate these consolidated GHG emissions to the relevant DIE (Case 4).

Cases 1 to 4 are can be applied to only 15% of known DIEs (either foreign-owned DIEs in France or French DIEs abroad), but those 15% actually account for 72% of French total FDI value, which appears as a more relevant share.

Case 5 arises when there is no listed entity in the entire chain of ownership. In this case, information is scarce, especially for those DIEs located abroad. In the absence of available microeconomic data, we experimented a mezzo-approach to estimate their GHG emissions. The idea behind was to approximate individual firm's GHG emissions (E_i) calculated as its value added multiplied by CO2 emission per unit of value added at country level, in those sectors of activity receiving the investment (1).

$$E_i = VA_i \times E_{s,c} \tag{1}$$

where i is the DIE unit, s, the sector f activity and c, the country

CO2 intensity figures being one of the UN Sustainable Development Goals' indicator, they are now quite widely available for many countries, including at a sectoral level. Estimating DIE's value added, however, required a leap of faith, albeit grounded on simple theory: production, hence value added, can roughly be considered as resulting from two main factors, capital and labour. Extensive academic work has focused on labour shares and there is data computed for many countries (see for example Guerriero, 2019). Based on the most basic production function, DIE's value added can be inferred from the DIE net results (the only available data) divided by the average capital share per sector/country (2).

$$VA_i = \frac{Net\ results_i}{1 - Labour\ Share_{s,c}} \Leftrightarrow \frac{Net\ results_i}{Capital\ Share_{s,c}} \tag{2}$$

This line of investigation proved to be a dead end. Labour shares at the sectoral level are only available for large industrialised countries, with a strong degree of heterogeneity in levels. This heterogeneity is driven by multiple factors: domestic labour market composition, policies, sectoral breakdown of the economy, etc. Eventually, our oversimplified model proved counterproductive and detrimental to our efforts as too much information needed to be assumed. Dealing with case 5 clearly illustrates a gaping business accounting data gap. Firms that fall into this category represent around 28% (in value) of FDI.

Conceptual framework and theoretical model to allocate GHG emissions within a multinational group

Disregarding Case 5, the following developments aim to establish a model applicable to cases 2 and 4 that allocates GHG emissions among the different entities within the same corporate group, whether subsidiaries or parent company. This allocation model is both simple and general, as it ensures the breakdown of the group's total GHG emissions across its various entities.

The working assumption of this model is as follows: we assume that GHG emissions — specifically for scopes 1 and 2 — are directly linked to the stock of tangible long-term assets. This assumption aligns with the definitions of scopes 1 and 2.

The model basically works in three steps: (i) group-level GHG emissions are distributed at the entity-level based on stocks of tangible assets; (ii) this initial estimation is adjusted to account for the average carbon intensity of the country/sector to which each entity belongs; (iii) second-round estimates are normalized so that final entity-level estimates are consistent with group-level carbon emissions.

Let us consider a corporate group composed of n units.

Let E represent the total emissions of the group, C_i the tangible long-term assets of unit i, and $\alpha_{c_i s_i y_i}$ the average carbon footprint of entities in sector of activity s, located in country c, during year y, where $i = 1, 2, \ldots, n$. To simplify the notation, since every variable is yearly, we do not index the equations by y.

The carbon footprint $\alpha_{c_i s_i}$, expressed in tCO2e/Million LCU per international dollar, is defined as the ratio between the total emissions of sector s in country c during year y, and the added value of sector s in country c during year y. The introduction of this variable allows for the allocation of the group's GHG emissions among each unit, accounting for the average environmental efficiency of sector/country/year.

We have the following equations:

- The initial allocation of emissions based on tangible long-term assets:

$$e_i = \frac{C_i}{\sum_{i=1}^{n} C_i} \times E \tag{3}$$

- The adjustment of emissions according to sector/country/year average environmental efficiency:

$$e_i^* = \alpha_{c_i s_i} \times e_i \tag{4}$$

- The normalization of values proportionally:

$$E_i = \frac{e_i^*}{\sum_{i=1}^{n} e_i^*} \times E \tag{5}$$

Thus, the sum of E_i is equal to the total emissions of the group, E. This model allows for the allocation of emissions among n units, taking into account their respective tangible long-term assets and sector/country/year average environmental efficiency, ensuring a complete breakdown of GHG emissions.

As previously mentioned, the literature addressing the relationship between FDI and carbon emissions offers ambiguous conclusions. The equivocal nature of the results can be attributed to two opposing effects: the 'pollution haven' effect and the 'pollution halo' effect.

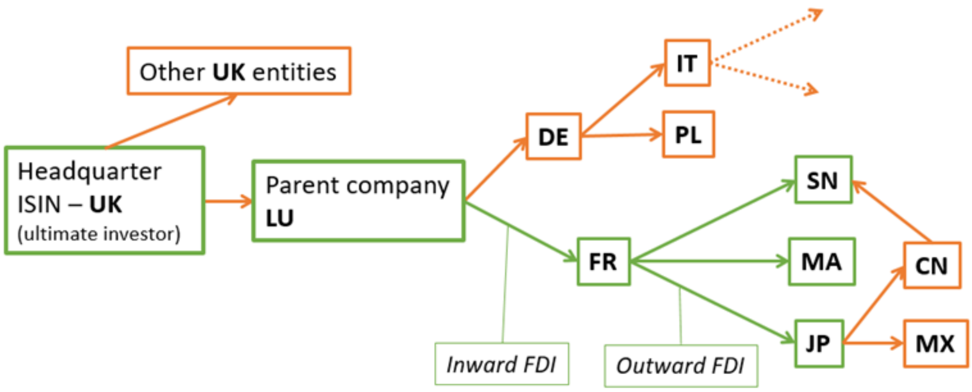
The initial allocation, in the first step, can be interpreted as assuming a pollution halo, as it presupposes identical production technology within the same productive structure. However, the second step, as it seeks to adjust the allocation to account for technological heterogeneity within the same productive structure, can be understood as assuming a pollution haven effect.

This model represents a first-best solution. In our case, data availability constraints are binding. Specifically, data on tangible long-term assets is largely available for listed groups, but much less for unlisted groups. Additionally, we do not have access to tangible long-term assets data for direct investment enterprises (DIEs). Employment data, that could be used as proxy, leads to the same setback.

Furthermore, the theoretical model assumes access to information on the entire productive structure of a group — meaning all affiliates must be accounted for. In our case, however, we only have FDI data on entities that are acquired or under the control or significant influence of French resident entities.

For inward FDI, we have information about French subsidiaries owned by a foreign parent company. That means we collect neither data about subsidiaries owned in other countries nor about entities in the investing country. For outward FDI, we only know about foreign subsidiaries owned by a French parent company (Figure 2).

Figure 2 – Example of the information French FDI data compilers have access to



Note: Domestic FDI compilation provides information in green, but data compilers remain unaware of all information in orange. In this fictive example, countries are written in ISO code. Dotted lines represent other possible FDI we do not know of.

Given the available information, the model outlined above needs to be adjusted to make it tractable.

A little detour via accounting is needed. In our case, at the DIE level, we have information on a portion of their equity, specifically the amount of the direct investor's shareholding, which represents the stock of direct investment in company i. By definition, equity includes tangible assets, as it helps finance long-term tangible assets such as the machinery of the invested companies. Therefore, in the process of full consolidation accounting, if we call A the parent company and B a subsidiary, the shares of B held by A are eliminated from A's balance sheet, and all assets and liabilities are added together line by line. In other words, the direct investor includes all the assets of the invested company in the assets side of its own balance sheet.

For listed groups for which we can retrieve carbon emissions data from the ISS database, we do have detailed balance sheet information, including long-term tangible assets and total assets, which we obtain from the Refinitiv database.

Therefore, by comparing the direct investor's shareholding in the DIE, i.e. the FDI, to the group's consolidated total of long-term tangible assets, we minimize (maximize) the contribution of the invested company to the group's long-term tangible assets if the direct investor's shareholding is lower (higher) than the invested company's long-term tangible assets.

Note that this approach works well for capital-intensive industries (e.g. manufacturing). However, this much less the case for labour-intensive industries (e.g. services), where total of long-term tangible assets can be smaller than the direct investor's shareholding in the DIE.

Hence we propose a variation that consists of comparing the direct investor's shareholding in the invested company to the group's consolidated total assets. In this case, by definition, the contribution of the invested company to the group's consolidated total assets will be lower than its actual contribution, as the invested company's equity is less than or equal to its total assets.

As a result, the aforementioned model will tend to underestimate GHG emission of the subsidiaries compared to their true values.

When the complete structure of a corporate group is unknown due to missing affiliates, it is essential to adjust the emissions allocation model to account for these gaps. Our approach introduces a fictitious "missing" affiliate to represent the collective impact of the unknown DIEs. This ensures that the total group GHG emissions are fully allocated, and the known DIEs are not unfairly overburdened with GHG emissions that should be attributed to missing units.

We then have the following equations:

- The initial allocation of emissions based on FDI and tangible long-term (total) assets:

$$e_i = \frac{FDI_i}{C_{group}} \times E \tag{6}$$

$$e_i = \frac{FDI_i}{TA_{group}} \times E \tag{6'}$$

  Where C_{group} is the group's tangible long-term assets and TA_{group}, the group's total assets, and FDI_i is the direct investor's share of capital in a DIE.

- The adjustment of emissions according to sector/country/year average environmental efficiency:

$$e_i^* = \alpha_{ics} \times e_i \tag{7}$$

- Unique fictitious missing affiliate compensating for unavailable information on group structure:

$$e_M^* = \frac{\left(C_{group} - \sum_{i=1}^{n} FDI_i\right)}{C_{group}} \times \overline{\alpha_{cs}} \times E, \qquad \overline{\alpha_{cs}} = \frac{1}{n}\sum_{i=1}^{n} \alpha_{ics} \tag{8}$$

- $$e_M^* = \frac{\left(TA_{group} - \sum_{i=1}^{n} FDI_i\right)}{TA_{group}} \times \overline{\alpha_{cs}} \times E, \ \overline{\alpha_{cs}} = \frac{1}{n}\sum_{i=1}^{n} \alpha_{ics} \tag{8'}$$

  The normalization of values proportionally:

$$E_i = \frac{e_i^*}{\sum_{i=1}^{n} e_i^* + e_M^*} \times E \tag{9}$$

Albeit not a panacea, the adjustment through a unique fictitious missing affiliate boils down to bunching together the contribution of GHG emissions by the parent company and all its potential unknown affiliates, including resident ones.

An alternative way to model a fictitious missing affiliate could be to simulate the average behaviour of the entities we know (mean imputation):

$$e_M^* = \frac{E}{TA_{group}} \frac{1}{n} \sum_{i=1}^{n} \alpha_{csi} \times FDI_i \tag{8''}$$

However, after computation, we noticed that the mean imputation for missing information (equation 8'') often leads to inconsistent results. This is due to the very partial view domestic FDI data compilers have on the global group structure of multinationals. For instance, in Figure 2 (see above), all the unknown entities in orange would be averaged into the fictitious subsidiary (depending on green entities' information). They would weigh as one single entity, while in reality they are six of them at least. As a result, this approach often leads to large overestimation of carbon footprint. Ideally, fostering more data sharing among FDI compilers could fix that issue. Hence we need to "compensate" by using equation (8) and consider the above model when we need to allocate GHG emissions within the group structure.

GHG emissions prediction model

For cases 1 and 2, GHG total emissions are available from listed companies – be it the DIE directly or the parent.

We use this piece of information to first build a simple GHG estimation model based on various variables, ranging from consolidated financial data to firms' characteristics. The data comes from ISS and Refinitiv. We develop a general linearized model (GLM) as GHG emissions data presents a lognormal distribution (see Table 1). For the following charts and tables, total GHG emissions correspond to the sum of scope 1 and scope 2 GHG emissions.

Table 1

Distribution of total GHG emissions, in CO2 tons equivalent

Listed entities providing the data (2016-2022)

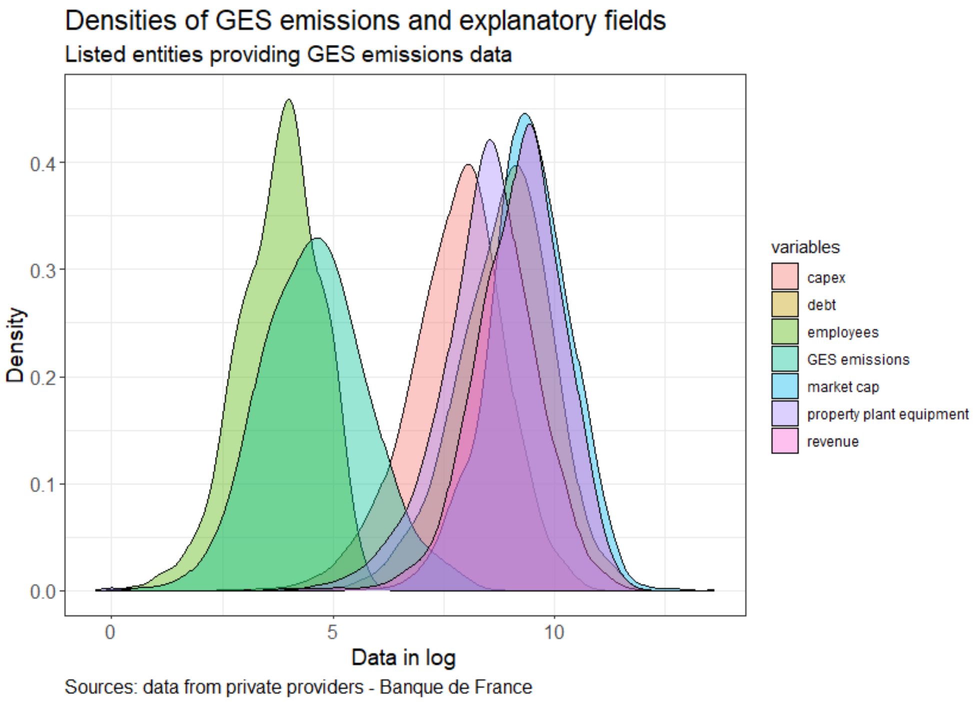
Statistics	Value
Minimum	4.4e-1
1st percentile	6.4e1
5th percentile	5.2e2
1st quartile	5.7e3
Median	3.7e4
3rd quartile	2.3e5

95th percentile	4.2e6
99th percentile	3.6e7
Maximum	5.2e8

Source: private data providers – Banque de France.
Note: median of total GHG emissions is 37,000 tons of CO2

Table 1 clearly shows the exponential shape of total GHG total emissions' distribution. A stepwise regression (using the Bayesian information criterion) helps us select the relevant numerical firm-specific determinants for GHG emissions reported in Figure 4.

Figure 4



Given the shape of their distribution, all these significant variables have been log-transformed. Their relation with GHG total emissions can be summarised into correlations (Table 2). Firms' financial situation is fairly closely correlated with GHG emissions with a higher debt and higher net revenues associated with higher GHG emissions. Even more striking, capital expenditure and physical equipment are highly correlated with GHG emissions, as well as firm size, proxied by the number of employees.

Table 2

Correlation of log-transformed numeric variables with total GHG emissions

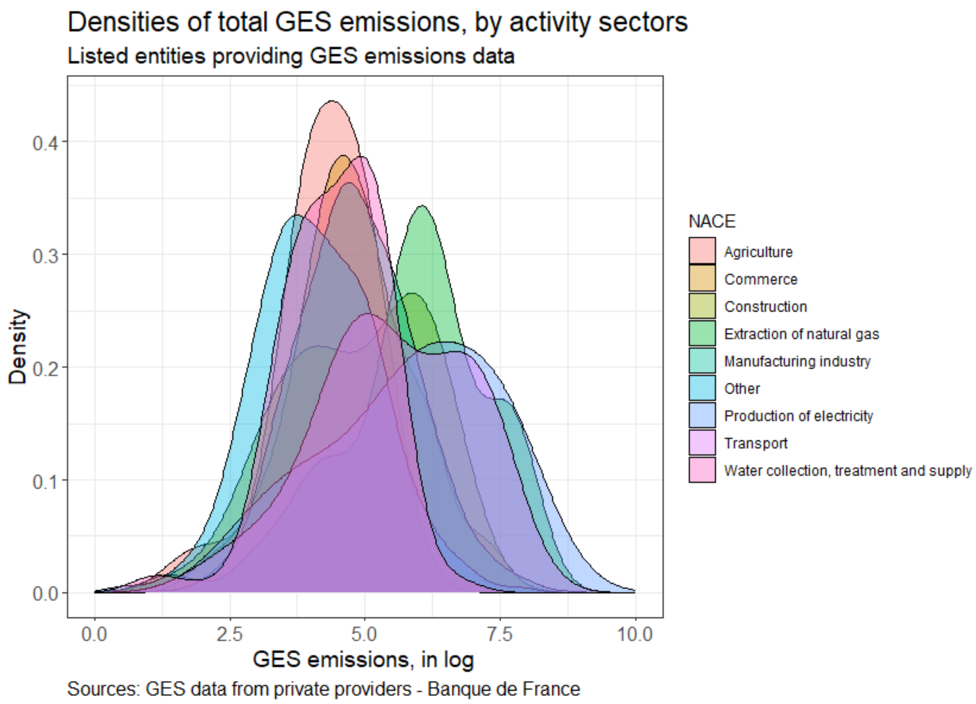
Listed entities providing the data (2016-2022)

Variable	Correlation value
Revenue	0.73
Employees	0.69
Market cap	0.58
Debt	0.63
Property plant equipment	0.79
Capex	0.75
Asset	0.59

Source: private data providers – Banque de France.
Note: variables about physical investments are the more correlated

Data is yearly from 2016 to 2022 and broken down by geographical zones (ex. Western Europe for France) and 18 wide sectors of activity, as GHG total emission also very much depend on the latter (see Figure 5).

Figure 5



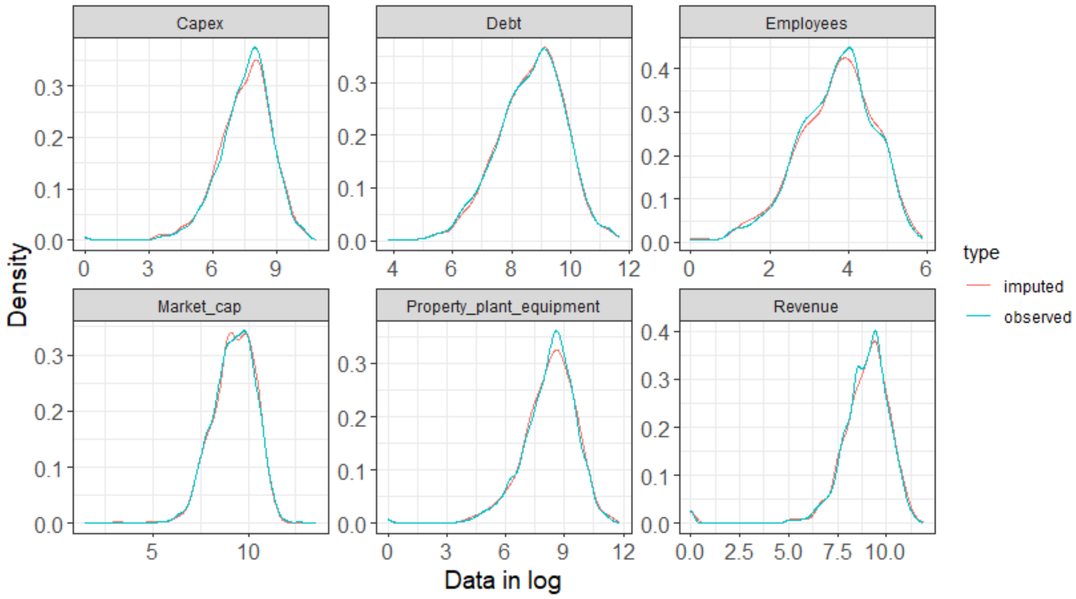
Based on this variable selection, the GLM equation may be written as follows, with X being the list of determinants selected above:

$$\log(GES_{total\_emissions}) = \beta X + \varepsilon$$

Numerical variables mostly come from Refinitiv and are very often properly completed for listed companies of our sample. However, we fill in missing observations using the Multivariate Imputation by Chained Equations (MICE) approach, similar to the nearest neighbours' based on available fields. Data imputation at this stage is believed to be benign as only 3 % of listed entities are missing one variable or more. Comparing variables' densities between imputed and observed data confirms the MICE imputation does not distort data too much (see Figure 6).

Figure 4

Densities of explanatory fields



Sources: data from private providers - Banque de France

Our GLM fitting (based on 1000 listed companies every year) provides satisfactory results, with a R^2 of 0.66 for our Scope 1 model and 0.60 for scope 2 model. When applying a 10-fold cross-validation, the out-of-sample mean fitting error reaches 2.2, which is deemed acceptable given the standard deviation of the modelled variables is 2.95. Most numerical variables appear significant (see Table in appendix). Firm size (proxied by the number of employees) appears as a significant positive determinant of GHG emissions : the larger the firm, the higher GHG emissions. Fixed equipment also appear significant and of the expected sign. Categorical variables appear overall less significant. Yet, significant sectors of activity to predict higher levels of GHG emissions is B (mining and quarrying) and D (which includes gas and electricity-related activities). Inversely, services sector predicts lower levels of emissions, particularly in sector K (financial and insurances) and J (information and communication), which appear sensible. Eastern Asia and sub-Saharan Africa are the only two very significant geographical determinants of GHG emissions, but the geographical zone matches the entity at a consolidated level, that may not well be the actual production region. Finally, the 2020 year dummy stands out, probably reflecting the downward impact of Covid on global GHG emissions.

In order to challenge these results, we run a Recursive Partitioning And Regression Tree (RPART). The results confirm those of our GLM model (see Table 3).

Table 3

Variable importance according to the RPART tree

Listed entities providing the data (2016-2022)

Variable	Relative Importance	
	Scope 1	Scope 2
Property plant equipment	18%	18%
Capex	19%	16%
Revenue	16%	19%
Debt	10%	10%
Employees	14%	22%
Market cap	11%	11%
Sector of activity	3%	1.5%

Source: Refinitiv – Banque de France.

Note: the tree was pruned before geo zones came out

The main conclusion from Table 3 is that fixed capital and revenue flows are indeed the more important variables to explain scope 1 GHG emissions, even at a consolidated level. Interestingly, when trying to estimate Scope 2 GHG emissions, firm size become the variable with most predictive power.

We can now use this GHG emissions model to estimated GHG emissions of listed DIEs (case 3) and of listed parent companies (case 4) that do not provide GHG total emissions.

2. Results

This data-intensive exercise provides an estimated carbon footprint measured in tons of emitted CO_2 for each DIE being part of the French FDI framework, outside the scope of Case 5 (i.e. for 28% of total French FDI in value). While this may seem a low coverage, our results look promising. By adding up DIEs' footprint depending on the FDI direction (outward or inward), we get two sets of data, reported in Figure 5 (for outward FDI) and 6 (for inward FDI).

Figure 5 pictures the carbon footprint of French FDI abroad in 2022. The darker the country, the higher GHG emissions due to French-owned DIEs. France has a higher density of resident multinational groups than other countries. In 2022, French FDI assets abroad stood at €1,414 billion, far above inward FDI, at only €746 billion. About half of French FDI abroad is located in other European countries, and clearly, the French carbon footprint is significant in neighbouring countries. Beyond Europe, the United States attract about 16% of French FDI and, in our sample, the French DIEs in the US emitted 674 900 tCO2e in 2022. Comparatively, China is only 2% of French outward FDI but the carbon footprint appears a lot worse, with more than 964 700 tCO2e emitted in 2022. This figure is likely to be largely underestimated since since China is one of the country that counts a lot of totally unlisted chains of ownership (Case 5).

Figure 6 shows the carbon footprint of foreign direct investors in France. Because investing countries are fewer, and mainly developed countries, there are fewer counterparties. FDI frameworks can be notoriously complicated to trace as multinationals have designed very complex corporate structure spanning over several countries and tax havens. Yet we note that the highest foreign carbon footprint in France is that of the United States, which is also the largest ultimate investor in France. Other European countries such as Germany, Spain or Italy are also top investors and their carbon footprint in France is significant.

Overall, the average carbon footprint of inward FDI is lower than that of outward FDI, which may be related to the CO2 emission efficiency of France, where energy production largely relies on low-carbon emitting nuclear power.

It is hard to benchmark our results, given we have only estimated a share of the French FDI carbon footprint. Based on a methodology close to that suggested by the IMF using input-ouput tables, a first estimate of the French carbon footprint of FDI stocks abroad had reached about 131 million tons of CO2 in 2014, about 0,5% of global emissions (Gigout, 2024) and amid a downward global trend in emission intensity. Our partial estimates are indeed below this benchmark.

Figure 5

Scopes 1 and 2 Carbon footprint of French direct investment abroad (tons of CO2; in logarithmic scale; 2022)

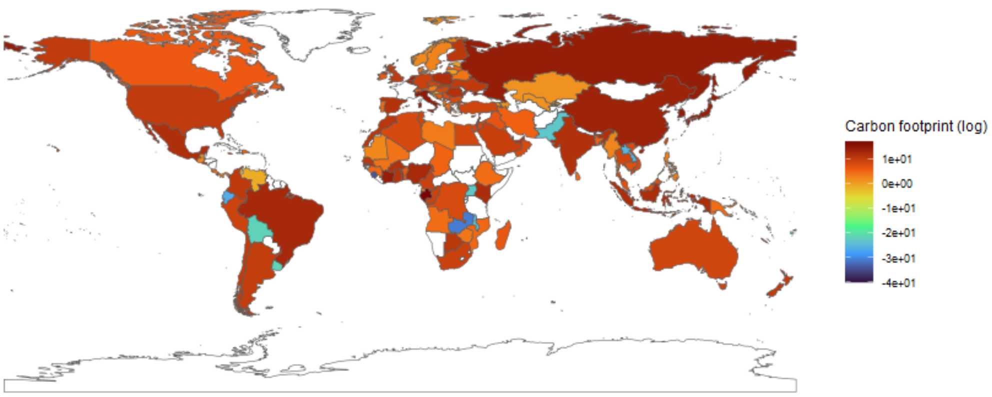
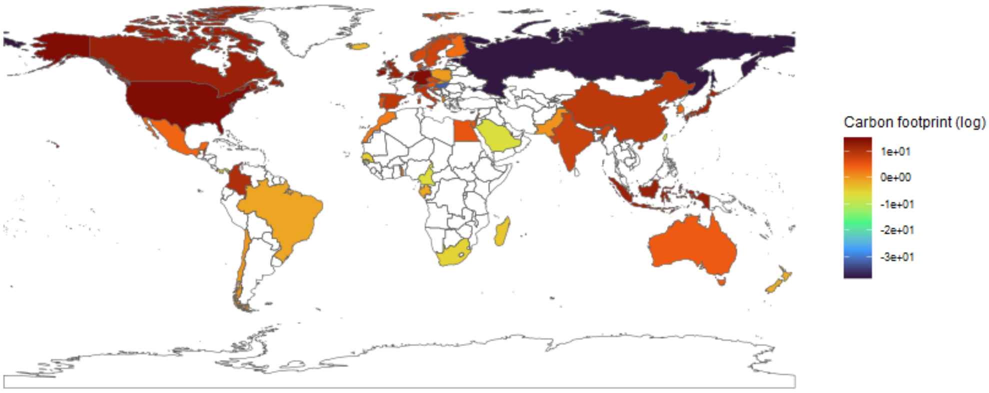


Figure 6

Scopes 1 and 2 Carbon footprint of foreign direct investment in France (tons of CO2; in logarithmic scale; 2022)



Source: Banque de France.

An interesting way to look at our results is to derive a CO2 intensity of FDI, by simply measuring GHG emissions divided by total FDI. Hence we can compare countries by looking at the environmental cost of one unit of FDI, whether outward or inward. Once again, our view is still partial since it is based on a fraction of total FDI.

Figure 7

Scopes 1 and 2 Carbon intensity of French direct investment abroad (in tons of CO2/FDI value; 2022)

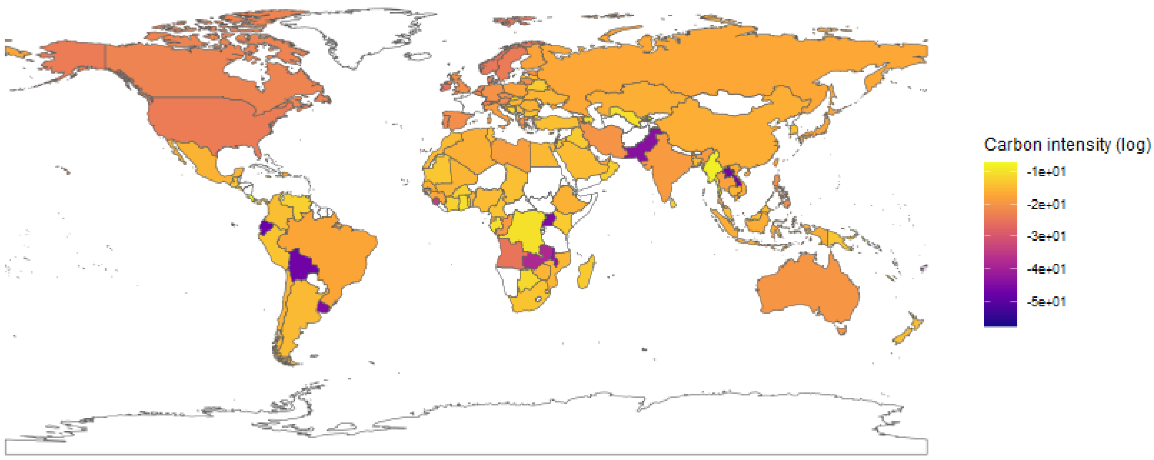
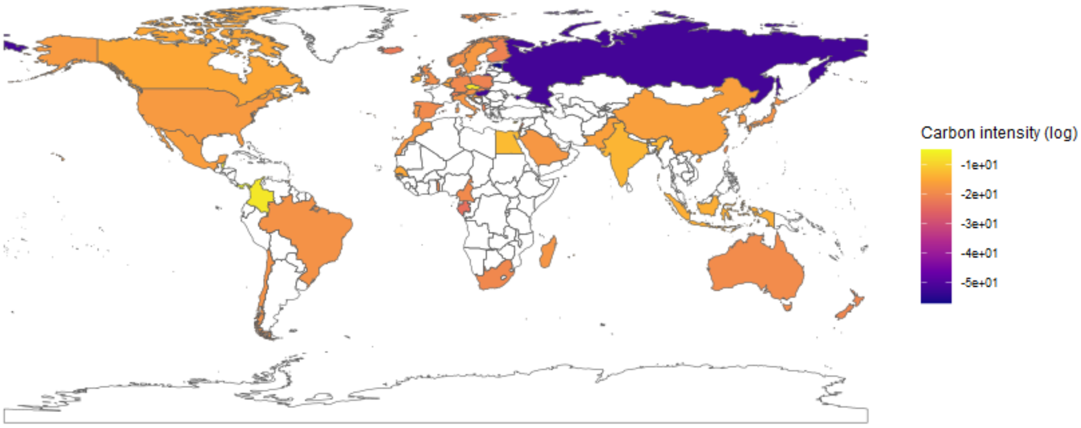


Figure 8

Scopes 1 and 2 Carbon intensity of foreign direct investment in France (in tons of CO2/FDI value; 2022)



Source: Banque de France.

3. Current limits and hopes

This paper presents a novel, data-driven approach to measuring the FDI carbon footprint using granular, firm-level data. This approach is close to FDI compilers' methods as they who often have to delve into companies' individual financial statements or investigate multinational group structures. By focusing on accumulated stocks rather than monetary flows, we also wanted to offer a more accurate reflection of how multinational enterprises contribute to GHG emissions through their global operations. While current approaches rely on macroeconomic methods and input-output tables, our micro-level approach provides a closer examination of emissions tied to individual firms, enhancing the precision and relevance of FDI carbon footprint assessments.

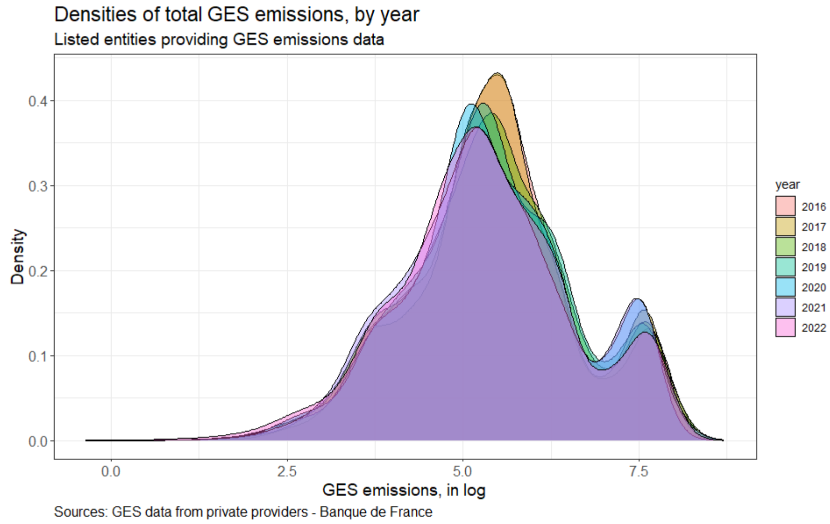
However, the results remain partial due to limitation in data availability. As our work progressed, a numerous amount of unforeseen questions and obstacles arose. First, we questioned the consistency and reliability of GHG emissions available both in corporate reports and estimated by private date providers. We starved for carbon efficiency data for less developed countries and for some information about the DIE's sector of activity, especially for outward FDI. Frustration peaked when trying to circumvent the lack of business accounting data for unlisted firms. Dealing with Case 5 unlisted firms was especially disappointing, although our mezzo-economic shortcut initially sounded promising. Yet, implausible results, in the absence of any "true" benchmark, become a quality measure. Despite a huge amount of work on data cleaning, consolidation and imputation, a considerable data gap still exists for almost a third of FDI assets. The GHG emissions allocation model is a step forward, but it is not without challenges either. While our use of a fictitious "missing affiliate" to address unknown entities in corporate groups is a promising technique, the way it is designed is not innocuous. For example, we tested an alternative model by which the fictitious DIE was granted the average emissions of known DIEs, which failed to provide credible results.

Despite these challenges, we are optimistic about useful developments in the short to medium term. Last year, several hundred non-financial entities have been testing a prototype climate indicator developed by Banque de France. The idea is to eventually add a new dimension to the Bank's activity of rating French firms' financial health. Business accounting data gap remains a clear sore spot. To run such a granular exercise, one needs more precise business accounting details about foreign DIEs. The new European Corporate Sustainability Reporting Directive (CSRD), applicable since January 2024, sets new standards and extra-financial reporting requirements for large companies and listed SMEs. Such regulatory advancements, together with the increasing availability of climate data, offer hope for more comprehensive and reliable assessments of FDI-related emissions. Over the past 15 years, a number of countries have established voluntary or mandatory GHG emissions measurement and reporting schemes with the aim to reduce global emissions. We are reaching the exciting stage where available data enables us to test different approaches suggested by international recommendations.

Looking forward, further research should focus on refining the micro-based methodology and comparing it with macro-level approaches to verify robustness. Improved business accounting practices that provides a more transparent view over global business structures of multinationals and more detailed emissions reporting will be crucial to fully understanding the carbon footprint of FDI.

Appendix

Yearly total GHG emissions (in log)



GLM fitting

GLM fitting for scope 1 GHG emissions

Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-2.244	0.246	-9.135	<2e-16	***
Year : 2017	0.030	0.036	0.844	0.399	
Year ; 2018	-0.009	0.036	-0.243	0.808	
Year : 2019	-0.022	0.035	-0.628	0.530	
Year : 2020	-0.120	0.035	-3.408	6.58e-4	***
Year : 2021	-0.063	0.036	-1.745	0.081	
Year : 2022	-0.013	0.039	-0.340	0.734	
Geo : Central Asia	0.187	0.396	0.471	0.637	
Geo : Eastern Asia	-0.018	0.171	-0.105	0.917	
Geo : Eastern Europe	0.026	0.197	0.131	0.895	
Geo : Latin America and the Caribbean	0.433	0.187	2.311	0.021	*
Geo : Northern Africa	0.059	0.191	0.311	0.756	
Geo : Northern America	0.048	0.1684	0.287	0.774	
Geo : Northern Europe	0.179	0.172	1.044	0.297	
Geo : South-eastern Asia	1.227	0.251	4.905	9.54e-07	***
Geo : Southern Asia	0.179	0.196	0.913	0.364	

Geo : Sub-Saharan Africa	0.731	0.203	3.605	3.14e-4	***
Geo : Western Asia	0.219	0.197	1.114	0.265	
Geo : Western Europe	0.123	0.168	0.735	0.462	
NAICS Sector : B	1.165	0.155	7.516	6.34e-14	***
NAICS Sector : C	0.087	0.135	0.644	0.519	
NAICS Sector : D	1.19	0.151	7.882	3.72e-15	***
NAICS Sector : E	-0.228	0.146	-1.554	0.120	
NAICS Sector : F	0.144	0.160	0.899	0.369	
NAICS Sector : G	-0.021	0.140	-0.150	0.880	
NAICS Sector : H	0.563	0.149	3.774	1.6e-4	***
NAICS Sector : I	-0.163	0.162	-1.005	0.315	
NAICS Sector : J	-0.788	0.138	-5.704	1.22e-08	***
NAICS Sector : K	-0.936	0.140	-6.687	2.46e-11	***
NAICS Sector : L	0.019	0.145	0.135	0.893	
NAICS Sector : M	-0.598	0.138	-4.332	1.50e-5	***
NAICS Sector : P	-0.117	0.423	-0.278	0.781	
NAICS Sector : Q	-0.145	0.202	-0.719	0.472	
NAICS Sector : R	-0.157	0.187	-0.840	0.401	
NAICS Sector : S	-0.586	0.232	-2.529	0.011	*
log(employees)	0.609	0.017	35.903	<2e-16	***
log(market cap)	-0.052	0.017	-3.111	1.08e-3	**
log(debt)	0.112	0.015	7.285	3.56e-13	***
log(property plan equipment)	0.334	0.018	18.287	<2e-16	***
log(capex)	0.098	0.019	5.075	3.98e-7	***

AIC : 16756
Number of Fischer Scoring Iterations : 2
R^2 : 0.66

GLM fitting for scope 2 GHG emissions

Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-1.191	0.231	-5.160	2.54e-7	***
Year : 2017	-0.062	0.033	-1.866	0.062	
Year ; 2018	-0.113	0.033	-3.405	0.001	***
Year : 2019	-0.172	0.033	-5.297	1.21e-7	***
Year : 2020	-0.278	0.033	-8.471	<2e-16	***
Year : 2021	-0.261	0.034	-7.760	9.71e-15	***
Year : 2022	-0.304	0.036	-8.429	<2e-16	***
Geo : Central Asia	0.195	0.369	0.529	0.597	
Geo : Eastern Asia	0.170	0.159	1.069	0.285	
Geo : Eastern Europe	0.108	0.183	0.590	0.555	
Geo : Latin America and the Caribbean	0.311	0.175	1.782	0.074	
Geo : Northern Africa	-0.087	0.178	-0.489	0.625	
Geo : Northern America	0.171	0.158	1.088	0.277	
Geo : Northern Europe	-0.011	0.160	-0.069	0.945	
Geo : South-eastern Asia	0.997	0.233	4.268	2.00e-5	***
Geo : Southern Asia	0.452	0.183	2.475	0.013	*
Geo : Sub-Saharan Africa	0.285	0.189	1.508	0.132	
Geo : Western Asia	0.151	0.183	0.823	0.411	
Geo : Western Europe	0.021	0.157	0.133	0.894	
NAICS Sector : B	0.794	0.144	5.492	4.1e-8	***
NAICS Sector : C	0.366	0.126	2.910	0.004	**
NAICS Sector : D	0.411	0.141	2.916	0.004	**
NAICS Sector : E	-0.207	0.137	-1.513	0.130	
NAICS Sector : F	-0.141	0.149	-0.946	0.344	
NAICS Sector : G	0.219	0.130	1.679	0.093	
NAICS Sector : H	-0.122	0.139	-0.873	0.383	
NAICS Sector : I	0.234	0.151	1.552	0.121	
NAICS Sector : J	-0.083	0.129	-0.643	0.520	
NAICS Sector : K	-0.365	0.130	-2.799	0.005	**
NAICS Sector : L	0.389	0.135	2.881	0.004	**
NAICS Sector : M	-0.152	0.129	-1.183	0.237	

NAICS Sector : P	-0.083	0.394	-0.210	0.834	
NAICS Sector : Q	0.332	0.188	1.762	0.078	
NAICS Sector : R	0.173	0.174	0.997	0.32	
NAICS Sector : S	-0.514	0.216	-2.381	0.012	*
log(employees)	0.578	0.016	34.961	<2e-16	***
log(revenue)	0.044	0.015	2.818	0.005	**
log(market cap)	-0.038	0.016	-2.366	0.019	*
log(debt)	0.068	0.014	4.788	1.72e-6	***
log(property plan equipment)	0.209	0.018	11.613	<2e-16	***
log(capex)	0.099	0.018	5.492	4.12e-8	***

AIC : 15777
Number of Fischer Scoring Iterations : 2
R^2 : 0.60

References

Borga M., Pegoue A., Legoff G., Sanchez Rodelgo A., Entaltsev D. and Gesga K. (2022), « Measuring carbon emissions of foreign direct investment in host economies », IMF Working Paper, WP/22/86, May.

Charpentier F., Martins B. and Bourcier P., (2023) « Estimating the carbon footprint of ICT using input-ouput analysis: dealing with overcounting and other challenges », 2023 International Conference on ICT for Sustainability (ICT4S) proceedings, pp. 154-163.

Gigout T. (2024), « Estimation de l'empreinte carbone des filiales françaises à l'étranger », Bulletin de la Banque de France, n°255, Article 1, November.

Gill F.L., Viswanathan K.K. and Karim M. Z. A. (2018), « The critical review of the pollution haven hypothesis », International Journal of Energy Economics and Policy, 8(1), 167-174.

Gopinath, M., & Chen, W. (2003). « Foreign direct investment and wages: a cross-country analysis », Journal of International Trade & Economic Development, 12(3), 285-309.

Guerriero, M. (2019), « The Labor Share of Income around the World: Evidence from a Panel Dataset Chapter 3 » in Fields, G. & Paul, S. (eds.) Labor Income Share in Asia: Conceptual Issues and the Drivers, Springer Singapore, Asian Development Bank Institute (ADBI) Series on Development Economics.

IMF (2008), « Balance of payments and international investment position manual », Sixth Edition (BPM6), Washington.

OECD (2009), « OECD Benchmark Definition of Foreign Direct Investment 2008 », Fourth Edition, Paris.

Robert J. Hodrick & Edward Prescott (1981), "Post-War U.S. Business Cycles: An Empirical Investigation," Discussion Papers 451, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

Sugiharti L., Zeqi Yasin M., Purwono R., Esquivias M.A., Pane D. (2022), « The FDI Spillover Effect on the Efficiency and Productivity of Manufacturing Firms: Its Implication on Open Innovation », Journal of Open Innovation: Technology, Market, and Complexity, 8(2), 99.

Tadesse, B., Shukralla, E. K. (2013), « The impact of foreign direct investment on horizontal export diversification: empirical evidence », Applied Economics, 45(2), 141–159.

Tukker, A.,de Koning, A., Owen, A., Lutter, S., Bruckner, M., Giljum S , Stadler K., Wood, R. and Hoekstra, R. (2018), « Towards robust, authoritative assessments of environmental impacts embodied in trade: Current state and recommendations », Journal of Industrial Ecology, 22 (3), pp. 585–598.

Tukker, A.,de Koning, A., Owen, A., Lutter, S., Bruckner, M., Giljum S , Stadler K., Wood, R. and Schmidt, S. (2020), « Towards accepted procedures for calculating international consumption-based carbon accounts », Climate Policy, 20 (sup1), S90–S106.

FDI CARBON FOOTPRINT: AN EXPERIMENT WITH GRANULAR DATA

VERONIQUE GENRE, ALICE MAGNIEZ, DAVID NEFZI & **FRANÇOIS ROBIN**

DEPUTY HEAD OF INTERNATIONAL TRADE AND FOREIGN DIRECT INVESTMENT DIVISION

BALANCE OF PAYMENTS DIRECTION

IFC, IZMIR – 07/05/2024

BANQUE DE FRANCE

EUROSYSTÈME

CONTEXT

Environmental impact of FDI unclear

Data Gaps Initiative 3 – Rec 3 = Carbon footprints of FDI

- IMF lead, with OECD

Methodology based on macro data (gross fixed capital formation) and input-output tables (cf. Borga et al., 2022)

- Could we do it the other round?

Combine firms' GHG reports with data compilers' FDI granular data

Comparison

Robustness checks
Data consistency
Reconcile top-down and bottom-up approaches

A GROWING AMOUNT OF GRANULAR GHG EMISSIONS' DATA

- **Private data providers for companies' GHG emissions**
  - Carbon 4 finance
  - Carbon Disclosure Project
  - Refinitiv
  - Institutional Shareholder Services
  - *Institut Louis Bachelier*

- Collect the data from annual reports or estimate it
- ~28,000 listed companies / year for Scopes 1 & 2

20%

80%

Robustness

- **Issue: listed companies only**
  - Listed French FDI in value:
    - **10% of assets**
    - **20% of liabilities**

OTHER AVAILABLE DATA

French FDI granular data
Banque de France

Direction	Siren (French ID)	Isin	Activity sector	Amount	Country
Asset	OK	*NA*	OK (investing)	OK (€)	
Liability	OK	*NA*	OK (invested)	OK (€)	

DIEs' parents data
Banque de France – INSEE (LIFI) – ECB (CSDB)

Direct parent	Isin	Greater parent	Country	Isin	*Case*
OK	OK	OK			*1-4*
OK	*NA*	OK		OK	*1-4*
OK	*NA*	OK		*NA*	*5*

GHG data
ISS – Refinitiv – C4F - CDP

Isin	Scope 1 & 2	Additional fields

Macro data
IMF – OECD – Banque de France

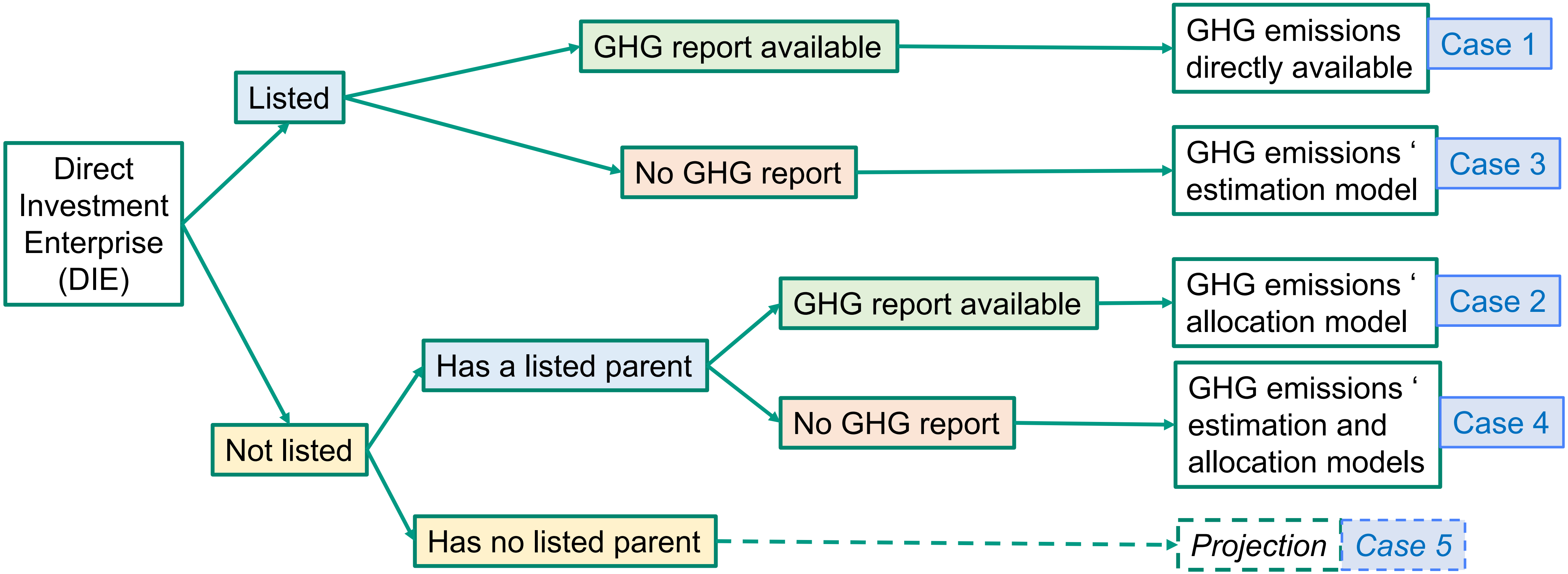
Country	Activity sector	Added value	PPP

Financial data
Refinitiv

Isin	NACE	Country	Debt	Capex	…

Yearly data 2016-2021

GHG EMISSIONS’ ESTIMATION FOR DIES



CASES IN MORE DETAILS

- **Cases 1 to 4**
  - **Estimation model**
    - Built from Cases 1 & 2 data

      6,000 ISIN

      $$GHG\ emissions \sim financial\ data$$

    - To be applied to Cases 3 & 4
  - **Allocation model for Cases 2 & 4**
    - What part of the DIE in the group total GHG emissions?

~49% in value

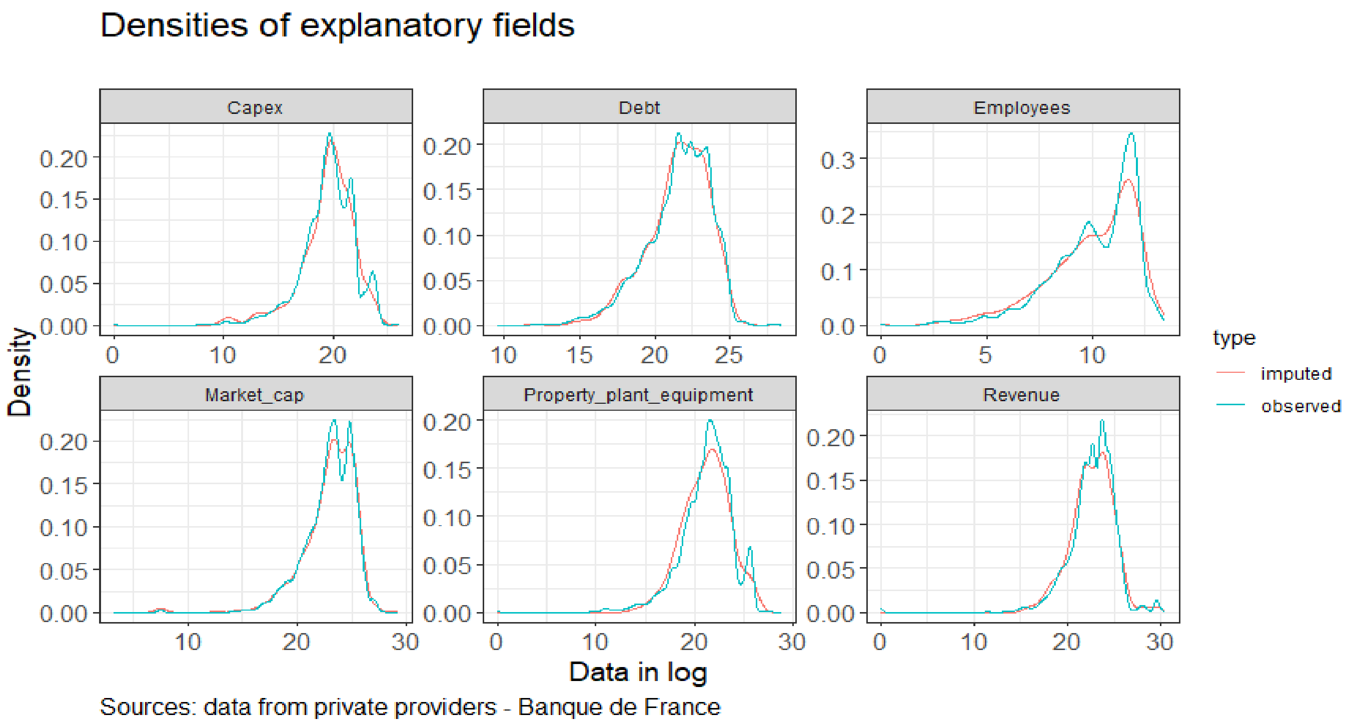
~1% in value

- **Case 5**
  - **Not dealt so far** (*it will in the next version of the paper*)
    - Idea: projection from Cases 1 to 4
    - Main challenge: data availability

~50% in value

GRANULAR DATA IMPUTATION

- **Some missing financial variables**
  - Multivariate Imputation by Chained Equations (MICE)



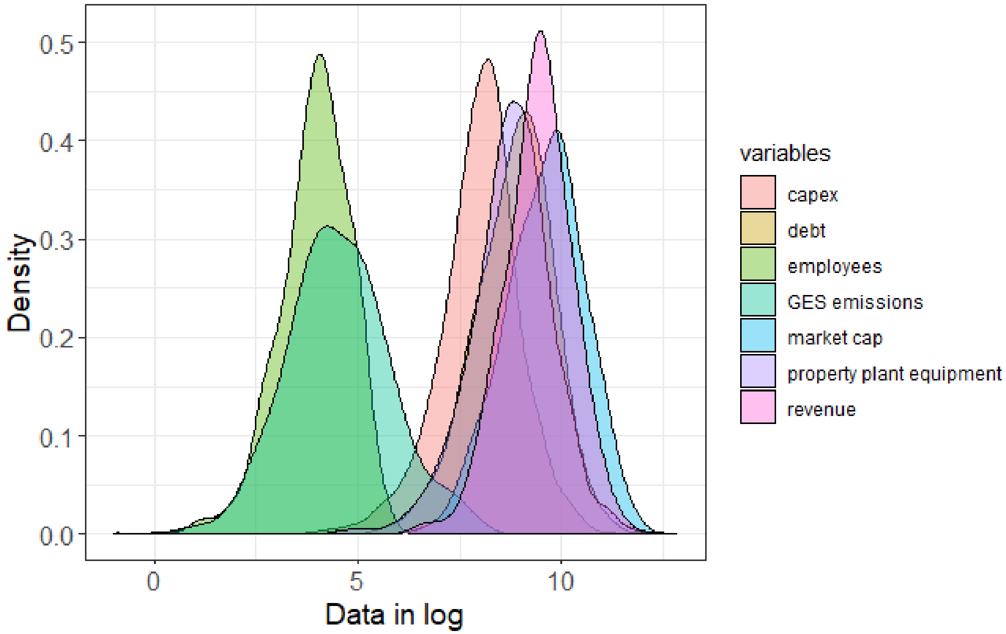
ESTIMATION MODEL

R²=0,7
Imputed data ~1% in value

- **Consolidated data for both financial and GHG emissions**
  - General Linearized Model (*might be improved later with boosting methods*)
  - Numeric variables and factors: activity sector and geography information

Densities of GES emissions and explanatory fields

Listed entities providing GES emissions data



Sources: data from private providers - Banque de France

Table 2

Correlation of log-transformed numeric variable with GHG total emissions

Listed entities providing the data (2016-2021)

Variable	Correlation value
Revenue	0.67
Employees	0.67
Market cap	0.56
Debt	0.59
Property plant equipment	0.75
Capex	0.71

Source: private data providers – Banque de France.

Note: variables about physical investments are the more correlated

ALLOCATION MODEL *(NB: we don't collect VA for DIEs)*

3rd best choice:
- Property plant equipment
- Employees (could be done for liabilities)

1) Holding ratio

$$GHG^{gross}_{DIE} = \frac{DIE_{\boxed{book\ value}}}{Parent_{long\ term\ investments}} GHG_{total\ emissions}$$

2) Correction step: from environmental efficiency

Energy efficiency

Missing data imputation:
1) Neighbour
2) Worst case

$$GHG^{corrected}_{DIE} = \boxed{\frac{GHG\ emissions_{country,sector}}{VA_{country,sector}}} GHG^{gross}_{DIE}$$

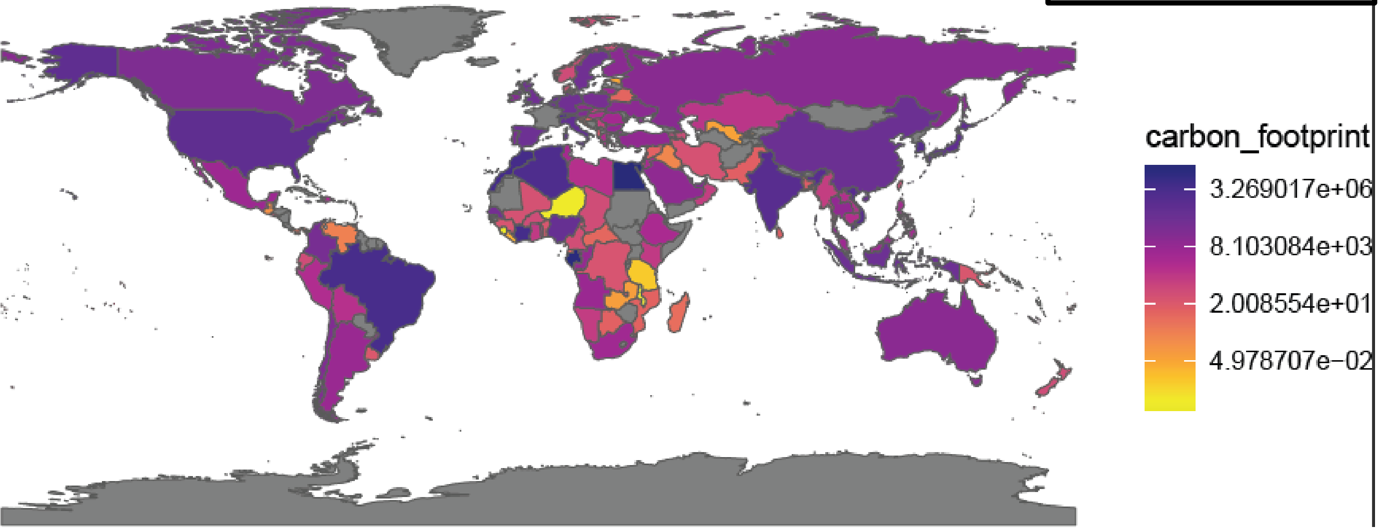
3) Normalization step (at a group level)

$$GHG_{Fictive\ DIE} = \overline{GHG^{corrected}_{Known\ DIEs}}$$

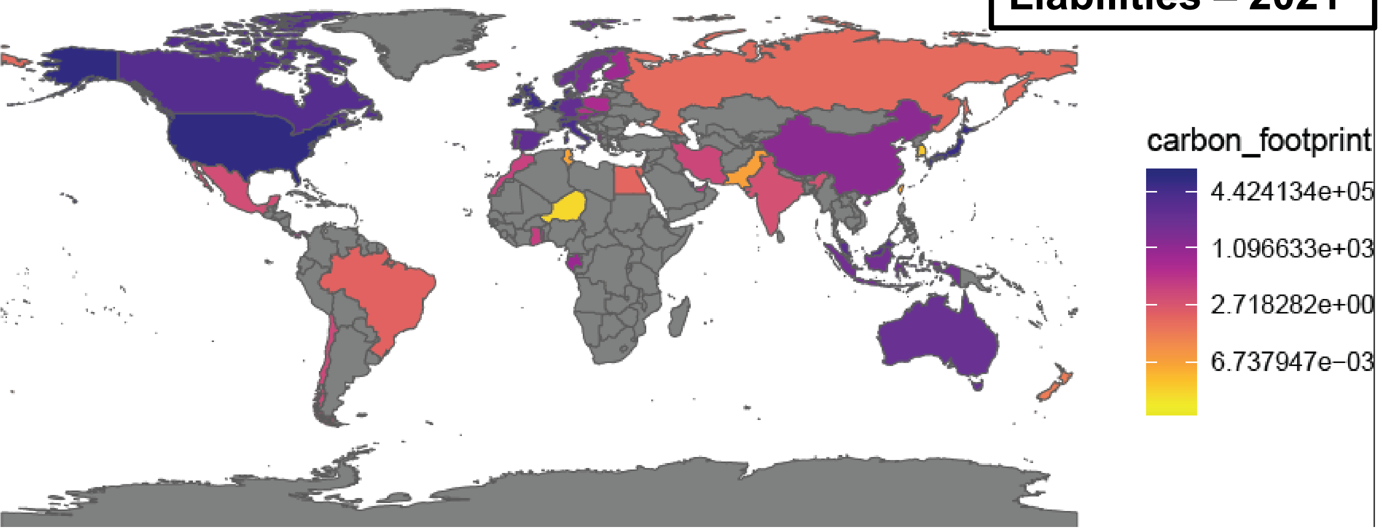
$$GHG_{DIE} = \frac{GHG^{corrected}_{DIE} GHG_{total\ emissions}}{GHG_{Fictive\ DIE} + GHG^{corrected}_{Known\ DIEs}}$$

RESULTS

Assets – 2021



Liabilities – 2021



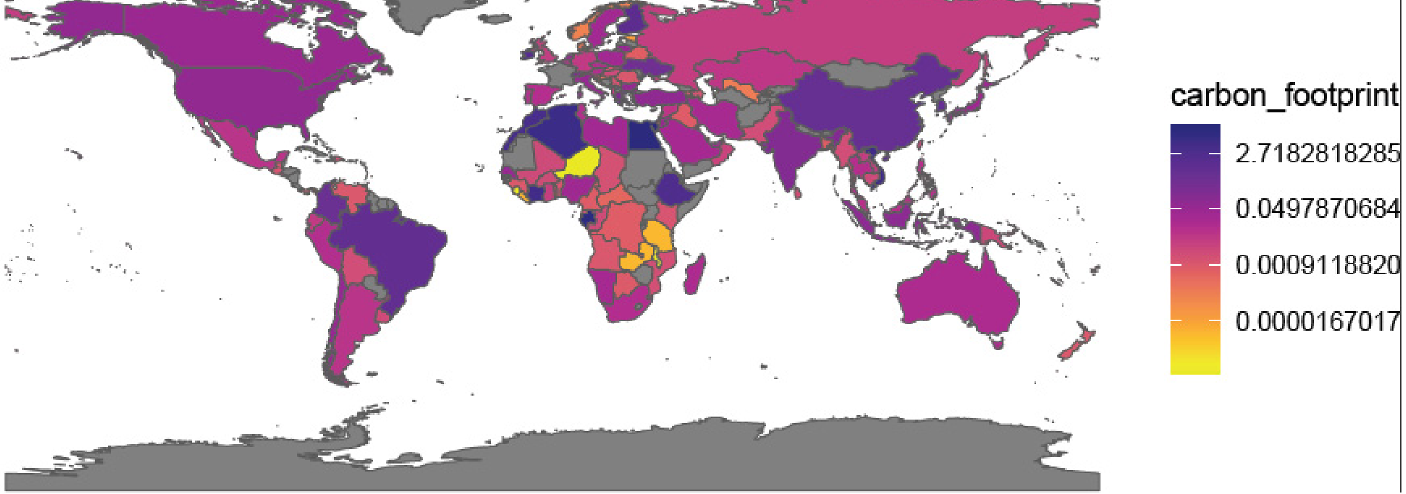
Log scale (tons of CO2)
~50% of FDI (in value)

Scope 1

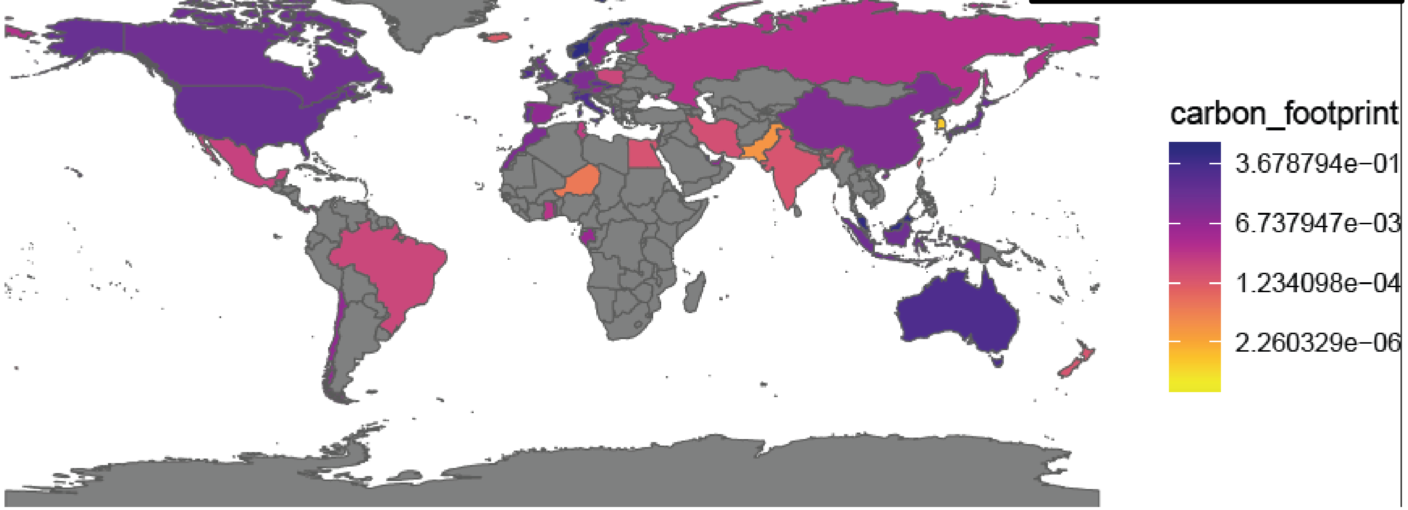
- Direct emissions

RESULTS – INTENSITY

Assets – 2021



Liabilities – 2021



Log scale (tons of CO2 / FDI value)
~50% of FDI (in value)

Scope 1 & 2

- Use the ratio of energy sector in the allocation model
- Some DIEs might own their provider
- Scope 1 of energy sector might duplicate with other sectors' Scope 2
- Better represents what GHG emissions DIEs are responsible of

CURRENT ISSUES DUE TO FDI COMPLEXITY

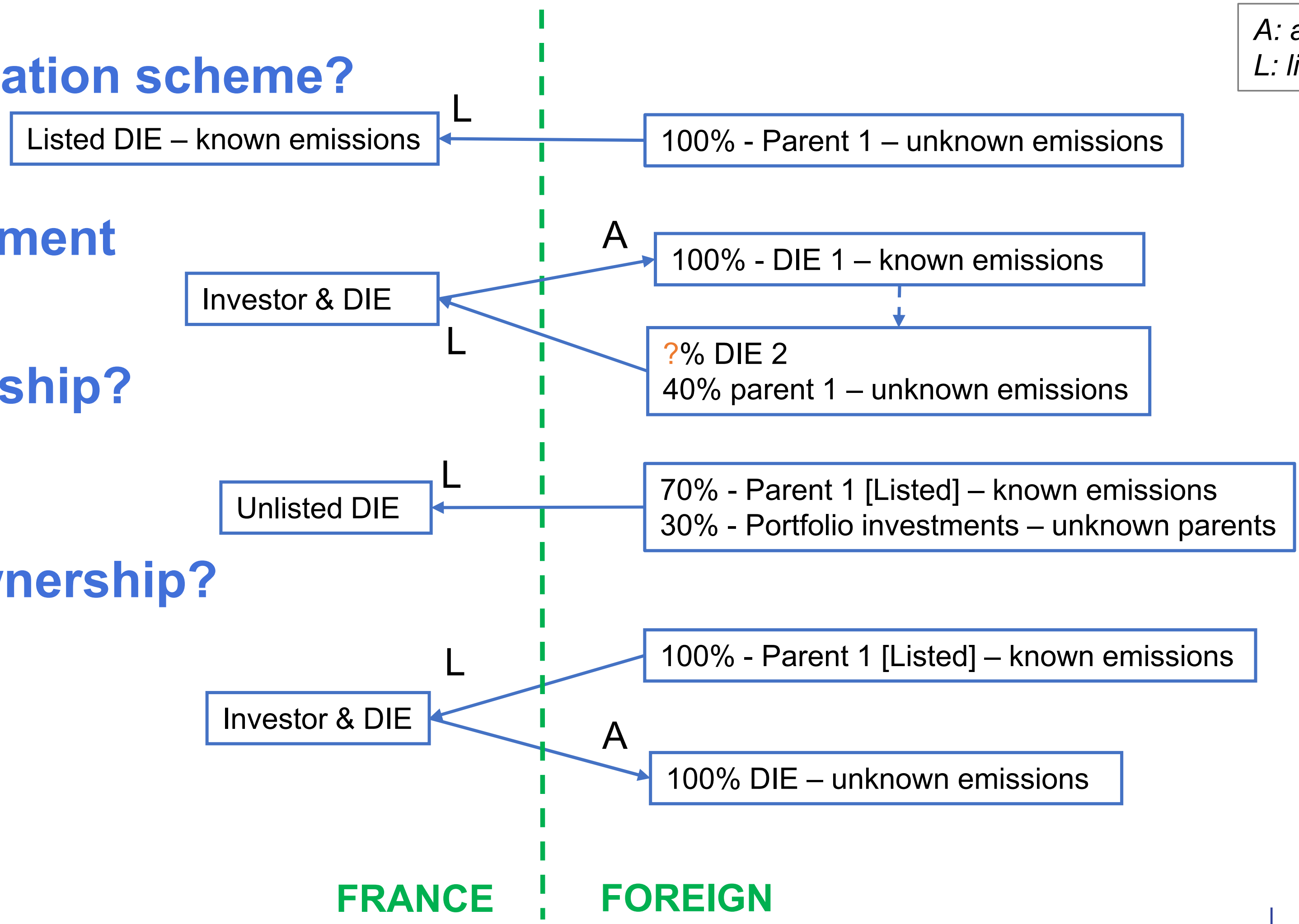
A: assets
L: liabilities

Invert the allocation scheme?

Circular investment

Multiple ownership?

Successive ownership?



LIMITS AND FUTURE WORK

- **Underestimates**
  - 50% FDI
- **Further improvements**
  - Explore other data for the imputation of energy efficiencies' ratios
  - Most of granular GHG emission data is estimated
- **Complementary approach**
  - Estimate value added (VA) at affiliates' level
  - Use VA to allocate emissions across DIE

$$GHG_{DIE} = VA_{DIE} \times \frac{GHG_{Country,Sector}}{VA_{Country,Sector}}$$

Some caveats will remain (VA estimation, successive ownership, missing data…)

BANQUE DE FRANCE

EUROSYSTÈME