Can we trust banks' emissions data?

Cristina Angelico* Enrico Bernardini*

April 2025

Abstract

High-quality banks' emissions data are key to assessing banks' climate-related transition risks and establishing reliable and quantifiable decarbonization goals. By exploiting a dataset sourced by four major data providers, we show that data coverage is often limited and information is highly heterogeneous across providers, especially for scope 3 emissions, which are the bulk of banks' carbon emissions. Scope 3 emissions are highly volatile over time, often lower than scope 2, and unrelated to the banks' exposure to high-emitting sectors. Finally, we show that Generative Artificial Intelligence-based scope 3 data are correlated with data collected from several data providers, therefore, they deliver useful information for checking data anomaly and consistency. However, they currently suffer the same limitations documented for the other sources, such as data quality and replicability. Regulatory developments are critical in fostering emissions disclosure of financial institutions and non-financial corporations in their portfolios, obtaining high-quality and reliable data, and enabling innovative techniques such as generative artificial intelligence.
Keywords: GHG emissions; Scope 3; Banks' climate-related risk

1 Introduction

There is a widespread awareness of the crucial importance of climate-related data to enable effective risk management and timely action, to tackle climate-related threats, and to pursue the ecological transition. Carbon emissions data are the focal climate metric, as testified by the growing regulatory initiatives aimed at moving forward from voluntary disclosure to mandatory reporting: in this vein, the provisions of the Corporate Sustainability Reporting Directive in the EU, the rules on climate disclosure by several US States,[1], the stock exchanges' obligations in mainland China, and globally almost 30 jurisdictions that have decided to adopt the Climate-related Disclosure Standard, issued in June 2023 by the International Sustainability Standard Board ISSB standards (IRFS 2023, IRFS 2024). Unlike ESG data and ratings (Berg et al. (2024)), the measurement of GHG emissions should have reached a standard thanks to GHG-Protocol (2013). However, there are still notable challenges regarding the consistency and quality of emissions data due to ambiguity in calculation methodologies and opaque representation, especially for the scope 3 emissions (i.e., indirect emissions related to upstream and downstream supply chain).

While previous research analyzed GHG data of non-financial firms, highlighting their limitations and the low correlation among data providers (Papadopoulos (2022); Bingler et al. (2022); Nguyen et al. (2023); Talbot & Boiral (2018); Busch et al. (2022)), this work seeks to shed light on the quality and consistency of GHG data of the European banking system, where banks play a central role in financial intermediation. We contribute to the existing literature by analyzing GHG data collected from multiple private data providers for a sample of large and medium listed banks in the Euro area

*Cristina Angelico and Enrico Bernardini, Bank of Italy, Climate Change and Sustainability Hub. The views expressed herein are those of the authors and do not necessarily represent the views of the Bank of Italy. All remaining errors are our own.
Contacts: cristina.angelico@bancaditalia.it and enrico.bernardini@bancaditalia.it.

[1]States of California, New York, New Jersey and Illinois are defining disclosure rules, while Security and Exchange Commission has eventually decided to withdraw its defense in a judicial litigation concerning the rules adopted on March 6, 2024.

to understand whether and how sourced and estimated data differ across providers, to what extent they are reliable and correlated to the banks' credit portfolio composition, and the challenges of using such data for research and supervisory purposes. Given the documented data gaps and limitations, we test whether emission data estimated with Generative Artificial Intelligence (GenAI) could be a novel and valuable source of information and comparison, and whether it has the potential to revolutionize research also in this area, beyond several others as discussed by Korinek (2023).

Financial institutions, including banks, face additional challenges in computing and handling their GHG emissions, compared to non-financial corporations. Indeed, beyond the emissions related to their operations (i.e., buildings, services or employees), they also need to assess emissions related to their credit and investment portfolios, ideally by gathering data concerning their non-financial counterparts (i.e., borrowers and security issuers) or otherwise exploiting sectoral emissions data as a proxy. This additional issue makes the case of GHG data for banks particularly interesting.

Furthermore, according to increasing regulatory requirements, banks are expected, first, to manage the risks deriving from the potential losses that an abrupt transition to a low-carbon economy could cause in their portfolios; second, to define portfolio decarbonization strategies as part of transition plans; third, to disclose their management framework and strategies on carbon emissions and climate-related risks. In this context, high-quality data on banks' counterparts carbon emissions are necessary for banks to assess their exposure to transition risk, define their decarbonization targets, and set their investment and credit policies. They are also necessary for investors, researchers, and policymakers to evaluate banks' climate-related risks and performance relative to their climate goals, both at the micro and macro levels.

Our analysis shows that GHG data for the Euro area banking system are still often incomplete, differ across providers, and display several unexpected and inexplicable features, especially for the scope 3 emissions, which, for banks, refer to emissions related to credit and investment portfolios. The documented data gaps and limitations jeopardize the reliability of scope 3 emissions and hinder their use for research and policy purposes.

Besides, while GenAI-based emission data might be able to complement the available information in the future, at the current stage, it cannot overcome the limitations documented for traditional data sources, as similar concerns arise mainly related to the data quality and replicability.

Our findings underscore the importance of improving the standardized disclosure of carbon emissions and estimation models, both for financial institutions and their non-financial counterparts, to enhance the quality and reliability of the data and enable cost-efficient innovative techniques such as generative artificial intelligence. Hopefully, the forthcoming EU Directive on disclosure (e.g., the Corporate Sustainability Reporting Directive), despite revisions under discussion as part of the Omnibus proposal, will help improve the coverage and consistency of carbon emission data.

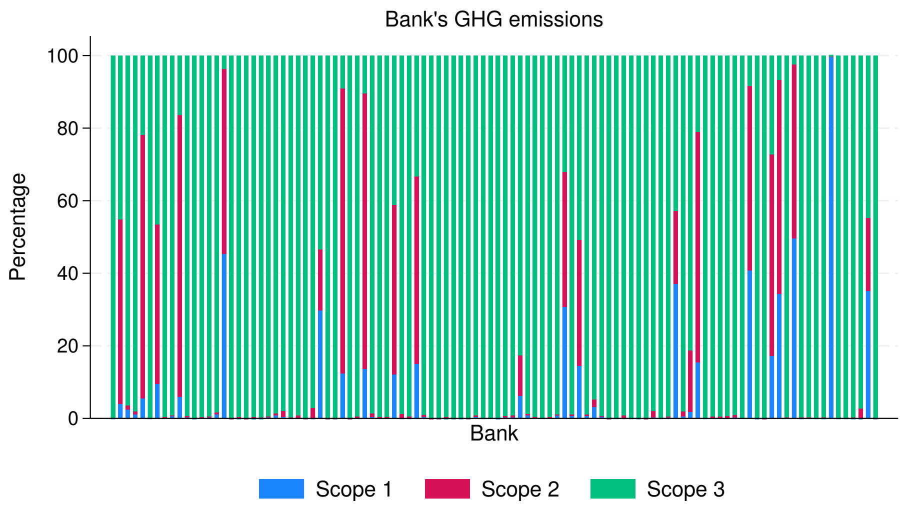
2 Data

We analyze different types of bank emissions according to the GHG Protocol classification, i.e. scope 1 (direct emissions under corporate control), scope 2 (indirect emissions, primarily related to the production of purchased electricity) and scope 3 which include all other indirect emissions in a company's value chain, divided into upstream and downstream sources. Upstream emissions derive from the production of input products or services (i.e., purchased goods and services, business travel, upstream transportation, and waste generation). In contrast, downstream emissions come from using and disposing of the products and services sold. For banks and other financial intermediaries, scope 3 downstream emissions are relative to loans and investment portfolios, i.e. financed emissions, classified as category 15 emissions: these are the emissions of the financed non-financial companies.

Most of the analysis focuses on the scope 3 emissions since they account for the largest share of the overall emissions of the banks in our sample, with an average share of 86% in 2022, which is even higher for most of the banks (Figure 1), although there are a few exceptions for which data are arguably misreported or poorly estimated.

The data sample covers 129 medium and large listed banks active in the Euro area between 2018 and 2023, whose emissions are sourced from four major private data providers: LSEG-Datastream, ISS, MSCI and Bloomberg. The methodology employed to compute the GHG data varies across providers. Usually, they rely on reported information published by credit institutions when available; otherwise, data are estimated using internal models or collected from other sources, such as the Climate Disclosure Project (CDP) survey, other reports, or websites.

Figure 1: Share of scope 1, 2 and 3 on overall emissions



Note: For each bank the figure plots the share of scope 1, 2 and 3 emissions over the total GHG bank's emissions. For each bank the emissions type data is the average over the different providers (Bloomberg, MSCI, ISS and LSEG-Datastream) as of the end of 2022. Each bar refers to a bank in the sample.

We combine GHG data with data of outstanding loans at bank and sector level from 2018 to 2023 sourced from Anacredit.[2] We focus on a few high-carbon sectors that will be key for the transition to a low-carbon economy according to the International Energy Agency (IEA (2023)); these sectors are manufacturing, energy, transport, and mining. Finally, we exploit data on scope 1, 2 and 3 emissions for our sample of banks in 2022 estimated with a GenAI tool named Claude, a large language model built by Anthropic.[3] Precisely, we run different requests using two models, Claude 3.5 and Claude 7, as described in Section 3.4. After a few iterations, Claude provided the requested data for all the banks in our sample. Data vary across different runs and include information classified as self-reported by banks in their sustainability reports and estimates based on their size and business models.

3 Results

3.1 Data coverage and heterogeneity across sources: aggregate and micro-level evidence

Data coverage on banks' emissions varies across metrics and sources. The coverage of the available GHG data for 2022 ranges between 41 and 121 banks depending on the scope, being more limited for scope 3 emissions (Figure 2, upper panel). This fact is surprising given that the sample includes

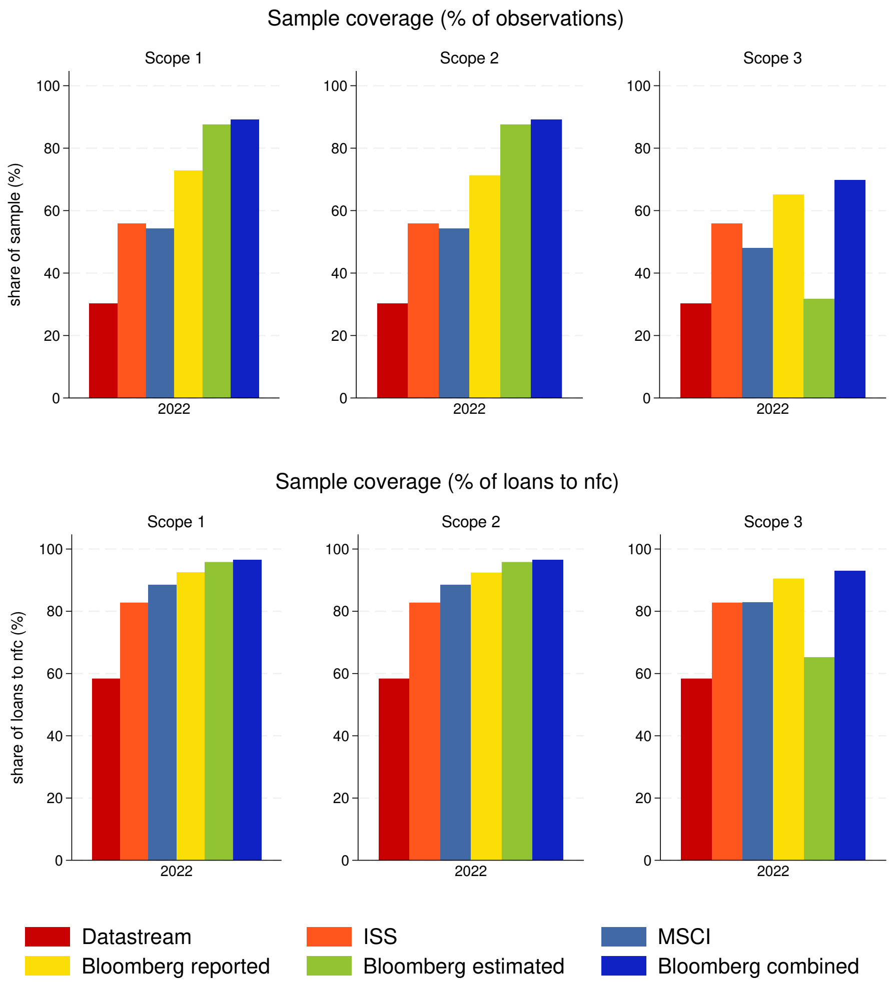
[2]AnaCredit is the Euro area credit register which collects at monthly frequency the amount and other details of any loan granted to corporations by Euro area credit institutions that exceeds the threshold of 25,000 euros.

[3]See *The Claude 3 Model Family: Opus, Sonnet, Haiku* (Anthropic (2024)) for more details.

only medium- and large-listed Euro area banks, many of which are already subject to disclosure requirements, while the others are expected to comply with such requirements in the next few years. Nevertheless, data are broadly available for the largest banks within the sample. The data available for 2022 cover a large share of loans to non-financial corporations granted by our bank sample: the coverage ratio ranges between 58% and almost 97% depending on the source and the scope of emissions (Figure 2, lower panel).[4]

At first glance, the average emission footprint of our bank sample varies substantially, in particular for scope 2 and 3 emissions, depending on the source (Figure 3, upper panel). Some heterogeneity observed across sources depends on differing coverage ratios, especially for scope 1 and 2, and on the outliers.

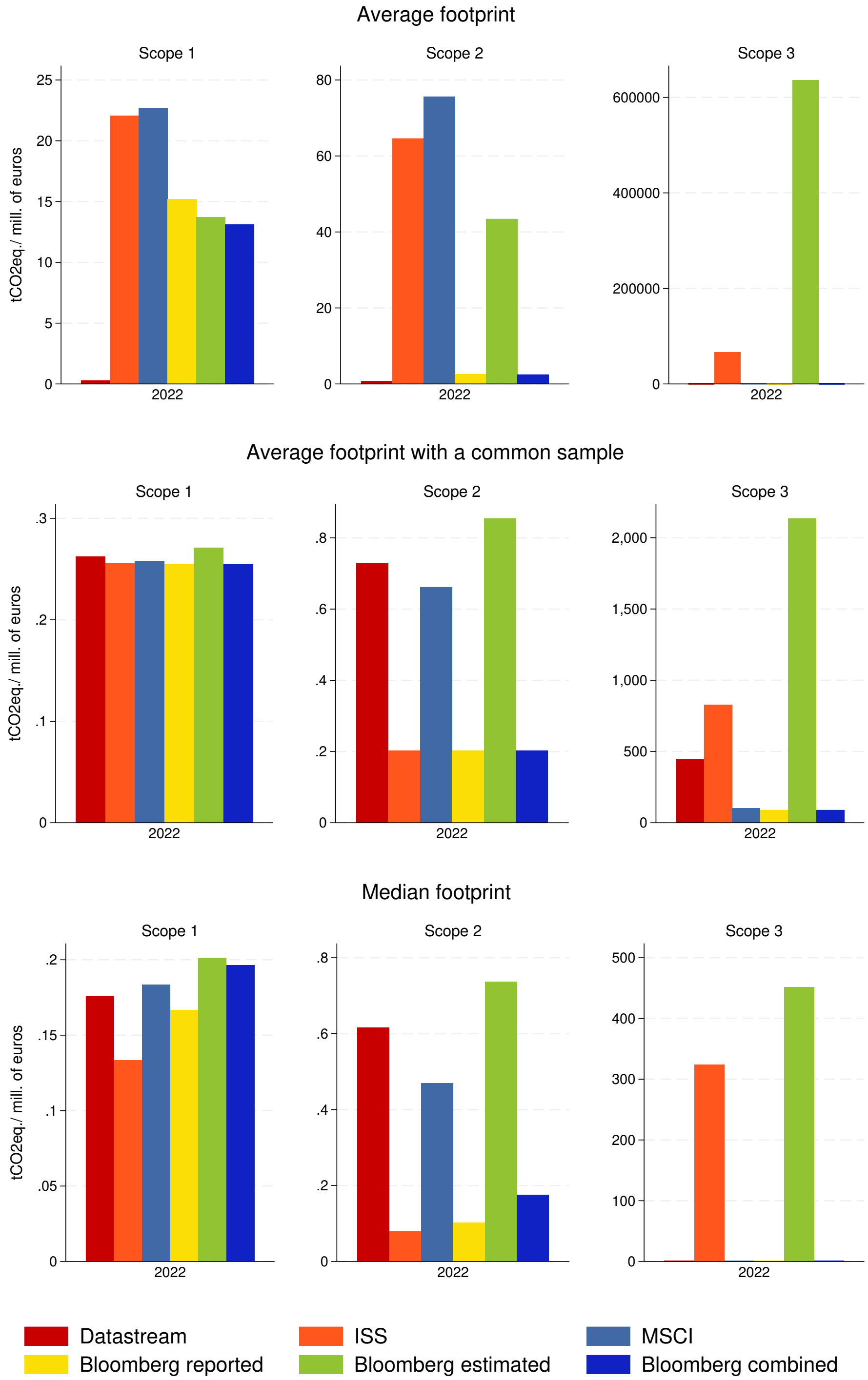
Figure 2: Sample data coverage for 2022 by source



Note: In the lower figure the sample data coverage is computed as the share of loans to non-financial corporations (nfc) in our sample covered by each data source as of the end of 2022.

[4] We focus on 2022 as we have more observations for this year; that may be because some providers still have not updated the 2023 data at the time we retrieved the data. The summary figures are mostly similar to those of previous years.

Figure 3: Emission footprint for the 2022 by source



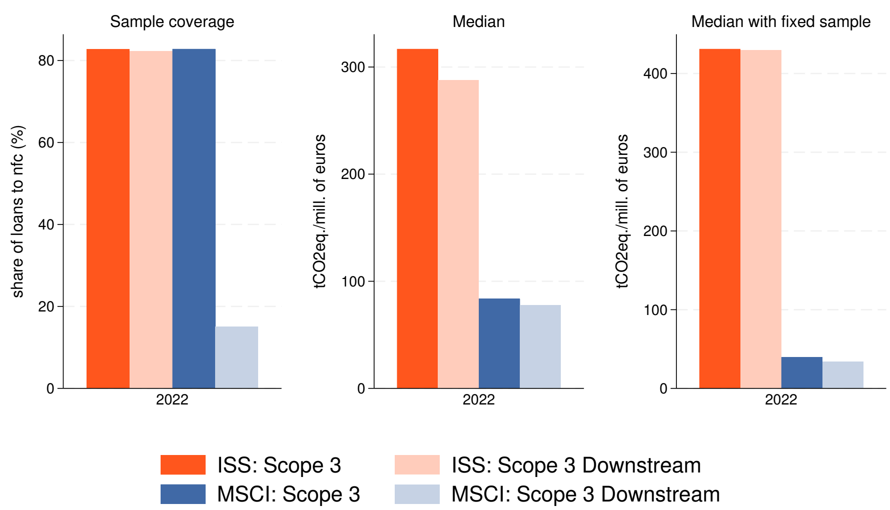
Note: For each bank, emission footprint is computed as the ratio between the bank's emissions and the total loans to non-financial corporations as of the end of 2022.

Indeed, once we restrict the sample to banks whose data are available from all four sources, the heterogeneity in average footprint substantially decreases, in particular for scope 1 emissions, which are widely reported (Figure 3, central panel). The heterogeneity also reduces for scope 2 emissions, although remarkable differences arise between reported and estimated emission footprint. A wide heterogeneity across data providers remains for scope 3 emissions. In this case, extreme values (arguably outliers) are observed for some providers, which rely mainly on estimated or self-reported data collected through the CDP questionnaire. As a matter of fact, it should be noted that the sample size for scope 3 emissions is substantially smaller, therefore less meaningful.[5]

The heterogeneity between sources is also lower when we look at the median rather than the average emission footprint (Figure 3, lower panel). In particular, the median values scope 3 emissions are much lower than the average. This is arguably due to the data distribution that presents a few outliers for almost all providers. A similar picture stands out considering only the sub-sample of banks for which we have data from all the data providers. According to the median scope 3 footprint, a loan of 1 million euros finances between 0.9 and 452 tons of CO2eq. The average footprint could instead reach higher values, above 600,000 tons of CO2eq. per million euros (Figure 3, upper panel). The average and median values are much larger when considering estimated data rather than reported emissions. The presence of outliers underscores the importance of looking at the median rather than the mean to obtain reasonable and comparable figures at bank level among alternative data sources, when possible, by checking the type of data (i.e., estimated or reported). At the same time, the fact that the average footprint is significantly higher than the median also emphasizes that, to have a comprehensive picture at system level, it is important to compute weighted average figures by including those banks with extreme values of carbon footprint (whenever consistent across providers), as they might represent a major source of climate risk within the financial system.

Downstream scope 3 emissions account for most of the total scope 3 emissions as they are related to the banks' portfolios (Figure 4).

Figure 4: Sample coverage and footprint for scope 3 and downstream emissions



Note: For each bank, emission footprint is computed as the ratio between the bank's emissions and the total loans to non-financial corporations as of the end of 2022. In the central panel we compare the overall median scope 3 emissions and the median scope 3 downstream emissions for a given source (the median is computed on a sub-sample for which we have both emission types for a given source). In the right-hand side panel, the median is computed on the sub-sample of banks for which we have the data from both providers.

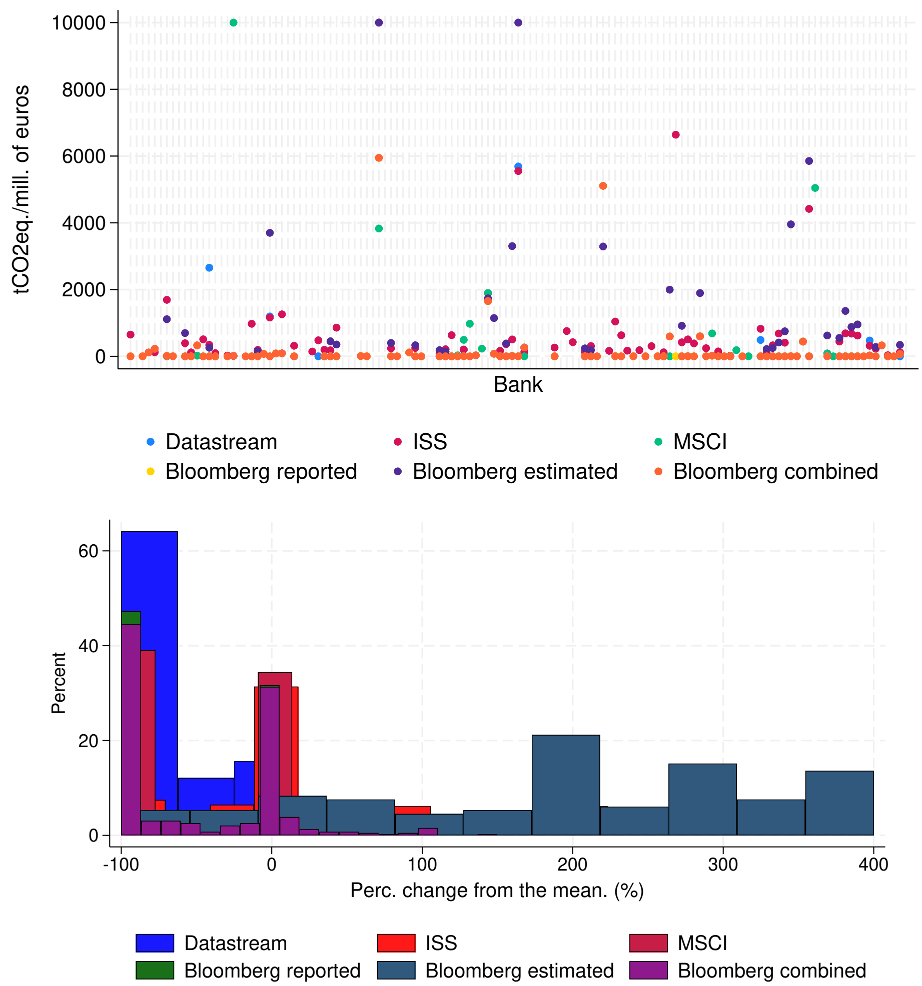
Despite their relevance, downstream scope 3 emissions are available only from two data providers

[5] We consider a sample of 40 banks for comparison of scope 1 and 2 emission footprint and only 14 banks for scope 3 emission footprint.

among those analyzed, and the coverage is minimal for one of those. Moreover, the two data sources disagree substantially in the median value, regardless the different sample coverage.[6]

More granular analysis based on bank-level emissions confirms a wide heterogeneity between data sources, especially for scope 3 emissions (Figure 5, upper panel) and scope 3 downstream emissions (Figure A1, in the Appendix 1).

Figure 5: Heterogeneity of scope 3 emissions across data sources



Note: Data on banks' GHG emissions by different providers as of 2022. In the upper plot, each dot refers to a bank in the sample. The figure was set with an upward bound for the emissions data for graphic purposes. On the x-axis, banks are sorted according to their loan size to non-financial corporations, so that the first bank on the left grants the smallest amount of loans. In the lower plot, the figure displays, for each data source, the histograms of the percentage change from the average scope 3 emissions computed as average across all the data sources for each bank and time. Data are bounded to be between -100% and +400%.

Several providers have high differences in the scope 3 emissions, with most values diverging from the average above 100%. At the same time, estimated data are significantly larger than the average

[6]The availability of on category 15 emissions data is even more limited (for both data sources), nevertheless when available category 15 and downstream scope 3 data are highly correlated.

in several instances, as well as providers that combine reported and estimated data tend to provide smaller estimates (Figure 5, lower panel).

Notwithstanding the above-mentioned heterogeneity, pairwise correlations of bank emissions from different sources between 2018 and 2022 are broadly positive and significant (Table 1). Consistent results are found for year-by-year correlations.

Table 1: Pairwise correlation across different data sources

	(1)	(2)	(3)	(4)	(5)	Average
			Scope 1 emissions			**0.881**
(2)	0.993***					
(3)	0.990***	0.995***				
(4)	0.949***	0.944***	0.997***			
(5)	0.945***	0.931***	0.325***	0.966***		
(6)	0.949***	0.930***	0.325***	0.998***	0.981***	
			Scope 2 emissions			**0.721**
(2)	0.840***					
(3)	0.919***	0.852***				
(4)	0.828***	0.566***	0.7782***			
(5)	0.926***	0.496***	0.340***	0.703***		
(6)	0.828***	0.5745***	0.456***	0.999***	0.712***	
			Scope 3 emissions			**0.569**
(2)	0.298***					
(3)	0.903***	0.306***				
(4)	0.906***	0.466***	0.842***			
(5)	0.521***	0.757***	-0.073	0.523***		
(6)	0.886***	0.462***	0.237***	0.974***	0.519***	
			Downstream scope 3 emissions			**0.994**
(3)		0.994***				

Note: Pairwise correlations using bank-level emission data from 2018 to 2022 per each data provider. In the table (1) refers to Datastream, (2) to ISS, (3) to MSCI, (4) to Bloomberg reported, (5) to Bloomberg estimated and (6) to Bloomberg combined.

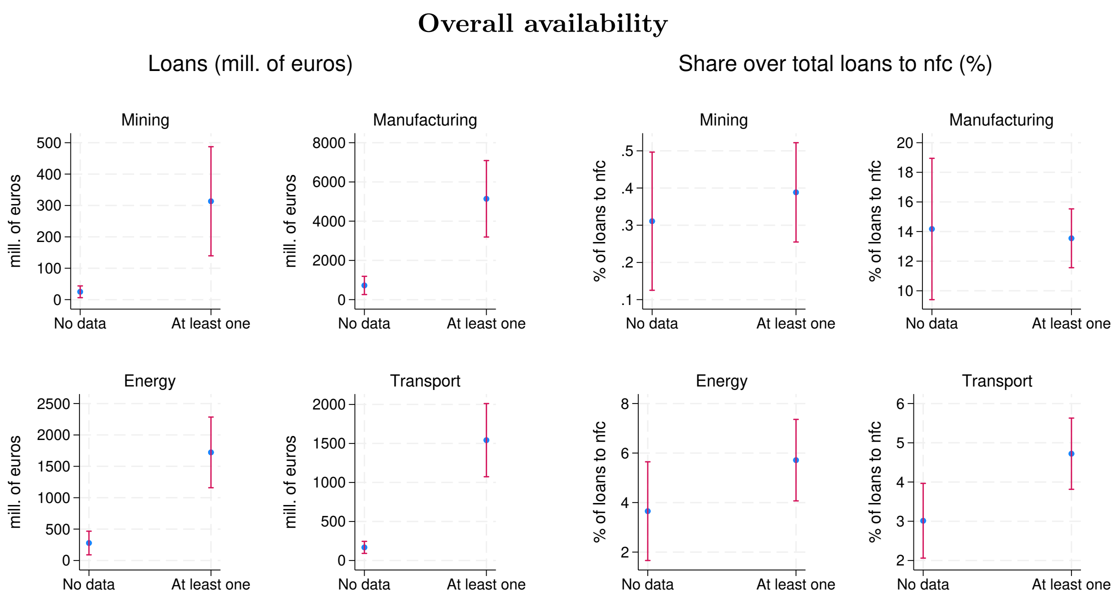
3.2 Banks' scope 3 emissions and exposure to high-carbon sectors

Within our sample, the availability of scope 3 emissions is positively correlated with the absolute size of bank loans rather than with their share to high-emitting sectors, such as mining (including oil and coal), energy, manufacturing (including iron, steel and chemicals) and transport (Figure 6). In other terms, emission data are more likely available for larger banks that grant more credit to high-emitting sectors, while the availability does not depend on the composition of the lending portfolio of banks. Indeed, the average share of loans granted to high-emitting sectors over the total loans to non-financial corporations is not significantly different between banks with available emission data and banks without available data. A similar picture emerges for self-reported scope 3 emissions according to Bloomberg which are more likely available for larger banks granting larger loans on average to high-emitting sectors (Figure A2 in the Appendix 1).

When available, scope 3 emissions are often positively correlated with the banks' portfolio size so that banks granting more credit to non-financial corporations have higher scope 3 emissions (Table 2, upper panel), while they do not show higher emissions footprints (middle panel). In particular, scope 3 emissions footprint does not reflect the banks' sectoral portfolio composition: the carbon footprints are surprisingly not higher for the banks more exposed to high-emitting sectors such as mining, energy and transport (Table 2, lower panel). However, banks that allocate a larger share of loans to the manufacturing sector display lower emissions footprints according to some data sources. This result may depend on the more miscellaneous nature of the sector.[7]

Similarly, the differences in banks' exposure to the high-emitting sectors do not explain the heterogeneity of scope 3 footprints between sources: the pairwise differences between the data sources are mainly uncorrelated with the share of loans granted to these sectors (Table A1 in the Appendix 1). A few exceptions stand out for mining, manufacturing, and transport. While this fact might be due to more disclosure of emission data by both banks and their non-financial borrowers, no clear pattern arise (Table A1, columns (1) and (10)). [8]

Figure 6: Scope 3 emissions availability and exposure to high-emitting sectors



Note: The figure displays the average (blue dot) amount of loans granted by the banks within our sample to high-emitting sectors (in millions of loans) and the average share of loans granted to such sectors over the total loans to non-financial corporations by a given bank in 2022. Confidence intervals are in red. Two groups are considered: 1) the banks for which we do not any scope 3 emissions data (No data); 2) those for which we have data on scope 3 emissions from at least one provider (At least one).

[7]Similar results are obtained if we include the total loans as a regressor or exclude the year fixed effect. Further, pairwise correlations suggest that emission footprints are often negatively correlated with the banks' exposure to such sectors.

[8]Nevertheless, the correlations are insignificant if we restrict the sample to 2022, possibly because of the low number of observations or an improvement in the quality of the data.

Table 2: Scope 3 emissions and footprint, banks size and credit portfolio

	(1)	(2)	(3)	(4)	(5)	(6)
	Scope 3 emissions					
Tot. loans	86.31	204.0***	-19.28	101.2	1009.9***	99.63
	(0.361)	(0.000)	(0.861)	(0.311)	(0.002)	(0.315)
R^2	0.084	0.449	0.001	0.089	0.491	0.082
Obs.	173	293	302	367	132	387
	Scope 3 emissions footprint					
Tot. loans	-0.00237	-0.375	-0.154	-0.189	-11.32	-0.183
	(0.176)	(0.317)	(0.265)	(0.315)	(0.332)	(0.316)
R^2	0.076	0.011	0.015	0.021	0.017	0.019
Obs.	173	293	302	367	132	387
	Scope 3 emissions footprint					
Mining	-158.6	-22846.7	-1900.9	-2.948	-903117.7	-0.861
	(0.165)	(0.295)	(0.416)	(0.959)	(0.315)	(0.987)
Manufacturing	-18.68*	-3214.6	-1321.0	-40.76*	-39226.9	-34.68*
	(0.053)	(0.249)	(0.273)	(0.077)	(0.400)	(0.083)
Energy	-41.51	-2354.8	-1435.5	-34.34	-348193.6	-30.35
	(0.307)	(0.374)	(0.319)	(0.155)	(0.323)	(0.163)
Transport	18.36	8433.6	-1147.0	79.61	298677.6	80.22
	(0.261)	(0.142)	(0.204)	(0.276)	(0.332)	(0.273)
R^2	0.132	0.040	0.027	0.119	0.089	0.106
Obs.	173	293	302	364	132	384

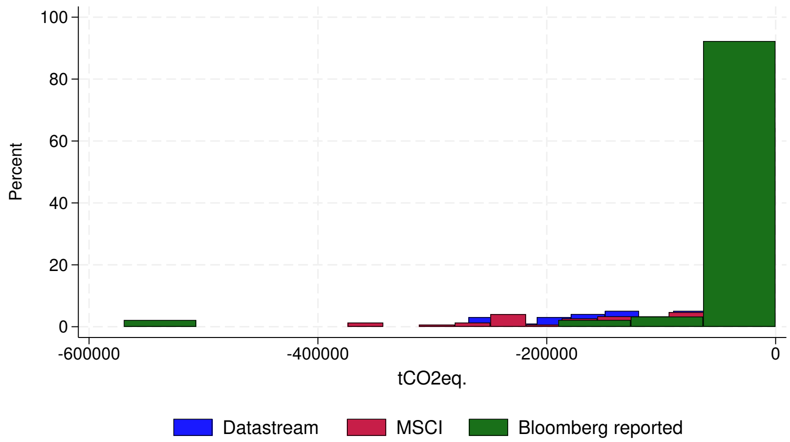
Note: Results from a set of regressions where the dependent variables are the bank's scope 3 emissions for a given data source (upper panel) or its scope 3 emissions footprint computed as the ratio between scope 3 emissions and the bank's total loans to non-financial corporations (middle and lower panels). *Tot. Loans* is the banks' total loans to non-financial corporations at a given year. In the lower panel, the independent variables are the share of bank loans to non-financial corporations granted to high-emitting sectors (mining, manufacturing, energy and transport). All the regressions include the constant and a year fixed effect. Each column refers to a data source: (1) is Datastream, (2) ISS, (3) MSCI, (4) Bloomberg reported, (5) Bloomberg estimated, (6) Bloomberg combined). Errors are clustered at the bank level. P-values in parentheses. * $= p < 0.10$, ** $= p < 0.05$, *** $= p < 0.01$

3.3 Inexplicable patterns of scope 3 emissions data

For some banks and several sources, scope 2 emissions are larger than scope 3. This occurs for about 50% of bank sub-sample covered by Datastream and MSCI and 27% covered by Bloomberg reported data. It never occurs with the data estimated by Bloomberg nor with the remaining sources. Among these anomalies in the data, the differences are significant in magnitude, with scope 3 emissions lower than scope 2 emissions by 60% on average (with a range between 1% and 100%, Figure 7). The relatively high frequency of this pattern suggests that it might be due to miss-reporting or underestimation of the scope 3 emissions rather than other reasons, such as the presence of banks managing large high-emitting data centres in our sample.

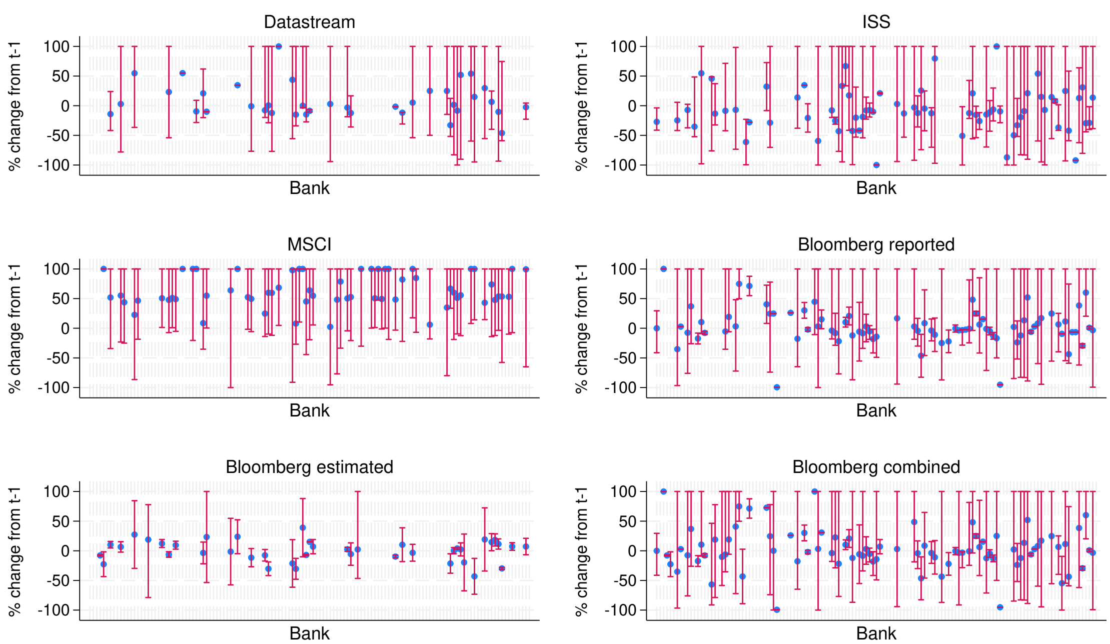
Scope 3 emissions data are also highly volatile over time. For all data sources, we observe that the banks' scope 3 absolute emissions and footprints vary substantially over the years, often increasing or decreasing by more than 100% (Figure 8). Such substantial changes occur over time per bank and source, with jumps and drops that often alternate. The variability might be explained by revisions in the methodology employed or in the emissions perimeter considered by reporting banks or providers when estimating the emissions. In particular, a methodological change could explain why, in a few years, the revisions by a given provider move in the same direction for all the banks.

Figure 7: Data anomalies: negative differences between scope 2 and 3



Note: The figures plots the histograms of the differences between scope 2 and 3 emissions when the differences are negative, by data source in 2022.

Figure 8: Data anomalies: bank-level annual percentage change in scope 3 emissions footprint

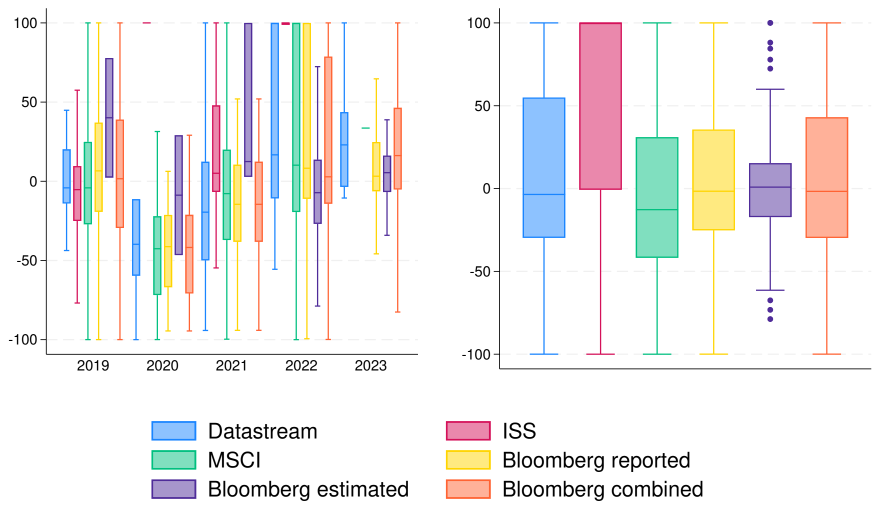


Note: The figures display the annual percentage change from the previous year in scope 3 emissions footprint reported for each bank by data sources. Each figure refers to a data source, while each dot/line refers to a bank. The blue dot is the median value, while the red bars display the minimum and the maximum. Outliers are not reported, and data are bounded to be between -100 and 100%. On the x-axis, banks are sorted according to their loans to non-financial corporations, so the first bank has the smallest total loans, while the last has the largest amount.

If we set lower and upper bounds for the data at -100% and 100% respectively, the average year-on-year percentage change of bank-level scope 3 emissions footprint ranges between 0.4 to 50%

depending on the data source (Figure 9).[9] However, a significant heterogeneity is observed within each source and over time.

Figure 9: Data anomalies: average percentage change in scope 3 emissions footprint



Note: The left-hand side figure displays box-plots of the average bank-level annual percentage change in scope 3 emissions footprint over time by data source. The right-hand side figure displays box-plots of the average bank-level annual percentage change in scope 3 emissions computed over the entire sample period by data source. Outliers are not reported, data are limited to be between -100 and 100%.

3.4 GenAI-based emissions data

Given the skyrocketing use of Generative Artificial Intelligence (GenAI) tools for an increasing range of applications, we test whether GenAI could be a novel and valuable alternative source of emissions data for the banking sector to fill the existing data gap and check anomalies. Previous works used Machine Learning techniques to estimate emissions (i.e., Assael et al. (2022); Nguyen et al. (2021)) or Large Language Models (LLM) to evaluate banks' environmental disclosure (Ángel Iván Moreno & Caminero (2023); Giannetti et al. (2023)) or to recover scope 3 emissions for non-financial corporations based on transaction data (Jain et al. (2023)).[10]

We asked Claude, a large language model built by Anthropic, to recover scope 1, 2 and 3 emissions data for the list of banks within our sample in 2022.[11] Precisely, to evaluate the robustness of the results across different runs and tools, we employed two different models developed by Anthropic: Claude 3.5 which is defined as the quickest model among those by Anthropic and adequate for most of the tasks, and Claude 7 which, instead, is defined as the most intelligent model and more suitable for mathematical tasks.

To our request, Claude 3.5 provided us with a table displaying the relevant data (hereafter, 1st run) by clarifying that the supplied data was not extracted from the banks' sustainability reports, but it was estimated to be representative of typical emissions for these institutions and should be verified

[9] In order to bound the percentage changes, we replace the values greater than 100 with 100 and those lower that -100 with -100.

[10] See Jain et al. (2023) for more information on other researches exploring the potential of LLMs in the domain of climate and sustainability. Bhuiyan (2024), instead, provides some case studies on how AI can help companies to measure and mitigate their own emissions.

[11] Prompt: *Can you please recover the scope 1, 2 and 3 emissions of the following banks in 2022? Can you please save the data in an excel file (that I can export) where there is a raw for each bank and 3 columns, one for scope 1, one for scope 2 and one for scope 3?*

against the banks' most recent official reports (see Appendix 2 4 for more details). Claude 7, instead, replied by providing a table that displayed : 1) reported data based on publicly available sustainability reports when available, and 2) estimated data, mostly for the smaller banks. Hereafter, we will refer to these data as 2nd run. The data provided by Claude presents several commonalities and similar challenges as those of other sources analyzed in the previous sections of the paper, including the mentioned anomalies in the ratio between the scope 2 and 3 emissions (in the 2nd run of the test). In the data provided by Claude 7 (2nd run), indeed, for about 20% of the banks the estimated scope 2 emissions were larger than scope 3. When inquired about these data anomalies, the tool recognized the mistake and corrected the data (hereafter, 3st run).[12] According to this novel data source and consistently with the others, scope 3 emissions footprint is, on average, much higher than the scope 1 and 2 emissions (Figure A3, upper panel, in the Appendix 2 4). Also in this case, the median values are substantially lower than the average footprints.[13]

By exploiting granular bank-level data two findings stand out. First, a positive correlation between bank emissions data provided in the different runs (Figure A4 in the Appendix 2), suggesting that similar approaches and consistent information sets are used despite the differences in the underlying LLM and the correction prompted during the runs (Table 3). Second, GenAI-based emissions are significantly correlated with the information provided by some private data providers (Table 3). In particular, the correlation of GenAI-based scope 3 emissions is stronger with the data provided by the sources that rely mainly or entirely on the estimated values (Table 3), suggesting the presence of some common driver in the estimation models and information set used by GenAI and data providers. The data classified as reported, instead, are different from and not significantly correlated with the data self-reported according to Bloomberg, possibly depending on the differences in the source and relevant reports exploited or due to miss-classification.[14] As for the other data sources, the scope 3 emissions footprints based on GenAI are not correlated with the size, nor with the portfolio composition (Table A2 in the Appendix 2).

Finally, we also asked Claude 7 to provide the financed (downward) emissions (category 15) for the banks in our sample. In a limited number of banks, it displayed estimated data computed as a percentage (ranging between 95 and 99%) of the total scope 3 emissions based on the standard portfolio size and composition for similar banks (Figure A5 in the Appendix 2).

In sum, the analysis suggests that at the current stage, GenAi-based emission data cannot overcome the limitations documented for the existing data sources. Although the data retain some information content since it is correlated with the data of the traditional sources, concerns related to the quality of the reported or estimated emissions arise, as for the other well-known sources analysed in the paper. Additional caution is also warranted on replicability of GenAi data, given the heterogeneity among data provided through subsequent runs.

[12] Claude clarified that the scope 3 emissions if correctly assessed should be substantially larger than the scope 2 as they relate to the financed emissions. The errors were mostly due to an underestimation of the scope 3 emissions for medium and smaller banks, for which only upward scope 3 emissions were available or considered.

[13] For the scope 1 and 2 emissions (especially in the 2nd and 3rd runs), the median is close to those obtained using the data supplied by private providers and discussed in Section 3.1.

[14] Further, according to Bloomberg self-reported data are available for 88 banks, while according to Claude only for 22.

Table 3: Correlation between GenAI-based emissions data and other data sources

	1st run	2nd run	2nd- est.	2nd- rep.	3rd run
			Scope 1 emissions		
2nd run	0.6295***				
2nd - est.	0.8427***	1.0000***			
2nd - rep.	0.5435*** .	1.0000***	.		
3rd run	0.6295***	1.0000***	1.0000***	1.0000***	
(1)	0.7446***	0.6173***	0.9725**	0.5964***	0.6173***
(2)	0.7736***	0.6944***	0.3686*	0.6455***	0.6944***
(3)	-0.0693	-0.0395	0.7605***	0.0443	-0.0395
(4)	-0.0322	-0.0156	0.5466***	-0.0559	-0.0156
(5)	-0.0007	0.0078	0.7091***	-0.0474	0.0078
(6)	-0.0163	-0.0052	0.7063***	-0.0559	-0.0052
			Scope 2 emissions		
2nd run	0.6156***				
2nd - est.	0.8059***	0.9574***			
2nd - rep.	0.6883***	0.5704***	.		
3rd run	0.6156***	1.0000***	0.9574***	0.5704***	
(1)	0.7804***	0.5774***	0.9251*	0.9367***	0.5774***
(2)	0.6371***	0.5366***	0.2162	0.6593***	0.6751***
(3)	0.0647	0.0115	0.8615***	0.4248***	0.0115
(4)	0.3141***	0.2731***	0.1555	0.3178**	0.2731***
(5)	0.2832***	0.2168**	0.3992***	0.5881***	0.2168**
(6)	0.2830***	0.2511***	0.6795***	0.3100**	0.2511***
			Scope 3 emissions		
2nd run	0.9067***				
2nd - est.	0.6539***	1.0000***			
2nd - rep.	0.8925***	1.0000***	.		
3rd run	0.9051***	0.9968***	0.8803***	0.9992***	
(1)	-0.0968	-0.0899	-0.1015	-0.1661	-0.1014
(2)	0.6667***	0.7330***	0.8002***	0.7179***	0.7199***
(3)	-0.0448	-0.0455	-0.1105	-0.0937	-0.0369
(4)	0.1433	0.3132***	-0.0351	0.2879	0.3122***
(5)	0.4645***	0.5507***	0.7155***	0.4291*	0.5393***
(6)	0.1455	0.3149***	-0.0321	0.2879	0.3139***

Note: Correlations using bank-level emission data as of 2022 per each data provider vis-à-vis GenAI-based emissions, under the different runs. *Est.* refers to estimated data, and Rep. to reported data according to Claude. In the table (1) refers to Datastream, (2) to ISS, (3) to MSCI, (4) to Bloomberg reported, (5) to Bloomberg estimated and (6) to Bloomberg combined.

4 Conclusions

This paper highlights that, despite the key role of banks' emission data for setting decarbonization goals by the banking system, there still exist a number of challenges affecting, in particular, scope 3 emissions that jeopardize their reliability, thus hindering their assessment by researchers and policymakers.

The paper highlights the heterogeneity among sources of scope 3 emissions and several data anomalies including the high volatility of such data over time, with frequent revisions, jumps and drops per each bank and provider as well as the missing correlation between scope 3 emissions footprints and banks' exposure to high-emitting sectors. These issues call for more transparency and stability in the estimation methodologies applied by credit institutions and data providers. To address these issues, it is of utmost importance to broaden the application of standard carbon measures and strengthen the climate-related disclosure requirements as envisaged by regulatory developments in the EU and other countries where ISSB standards will be adopted. Last, the test of an innovative application of a GenAI tool for estimating carbon emissions suggests that, although it might be a promising technique to fill the data gap in the future, at the current stage, concerns related to replicability and data quality arise, calling for further investigation.

In summary, these findings corroborate the importance of improving the standardized disclosure on carbon emissions and estimation models, both for financial institutions and the non-financial corporations included in their portfolios. In this direction, regulatory developments play a critical role also to enable cost-efficient innovative techniques such as generative artificial intelligence.

References

Anthropic (2024), The claude 3 model family: Opus, sonnet, haiku.
URL: *https://api.semanticscholar.org/CorpusID:268232499*

Assael, J., Heurtebize, T., Carlier, L. & Soupé, F. (2022), 'Greenhouse gases emissions: estimating corporate non-reported emissions using interpretable machine learning'.
URL: *https://arxiv.org/abs/2212.10844*

Berg, F., Huidobro, O. J. & Rigobon, R. (2024), 'On the importance of assurance in carbon accounting'.
URL: *https://papers.ssrn.com/sol3/papers.cfm?abstract*$_i d = 4734240$

Bhuiyan, M. (2024), Carbon footprint measurement and mitigation using AI, Technical report, Available at SSRN: https://ssrn.com/abstract=4746446 or http://dx.doi.org/10.2139/ssrn.4746446.

Bingler, J. A., Colesanti Senni, C. & Monnin, P. (2022), 'Understand what you measure: Where climate transition risk metrics converge and why they diverge', *Finance Research Letters* **50**, 103265.
URL: *https://www.sciencedirect.com/science/article/pii/S1544612322004561*

Busch, T., Johnson, M. & Pioch, T. (2022), 'Corporate carbon performance data: Quo vadis?', *Journal of Industrial Ecology* **26**(1), 350–363.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/jiec.13008*

GHG-Protocol (2013), Technical guidance for calculating scope 3 emissions, Technical report.

Giannetti, M., Jasova, M., Loumioti, M. & Mendicino, C. (2023), 'glossy green' banks: The disconnect between environmental disclosures and lending activities, Technical Report 919/2023, Swedish House of Finance Research Paper.
URL: *https://ssrn.com/abstract=4424081*

IEA (2023), Net zero roadmap: A global pathway to keep the 1.5 °c goal in reach, Technical report.

IRFS (2023), Sustainability disclosure standard, climate-related disclosures, Technical report.
URL: *https://www.ifrs. org/content/dam/ifrs/publications/pdf-standards-issb/english/2023/issued/part-a/issb-2023-a-ifrs-s2-climate-relateddisclosures.pdf?bypass=on.*

IRFS (2024), Progress on corporate climate-related disclosures—2024 report, Technical report.

Jain, A., Padmanaban, M., Hazra, J., Godbole, S. & Weldemariam, K. (2023), 'Supply chain emission estimation using large language models'.
URL: *https://arxiv.org/abs/2308.01741*

Korinek, A. (2023), ‘Generative ai for economic research: Use cases and implications for economists’, *Journal of Economic Literature* **61**(4), 1281–1317.
URL: *https://www.aeaweb.org/articles?id=10.1257/jel.20231736*

Nguyen, Q., Diaz-Rainey, I., Kitto, A., McNeil, B. I., Pittman, N. A. & Zhang, R. (2023), ‘Scope 3 emissions: Data quality and machine learning prediction accuracy’, *PLOS Climate* **2**(11), 1–31.
URL: *https://doi.org/10.1371/journal.pclm.0000208*

Nguyen, Q., Diaz-Rainey, I. & Kuruppuarachchi, D. (2021), ‘Predicting corporate carbon footprints for climate finance risk analyses: A machine learning approach’, *Energy Economics* **95**(C), S0140988321000347.
URL: *https://EconPapers.repec.org/RePEc:eee:eneeco:v:95:y:2021:i:c:s0140988321000347*

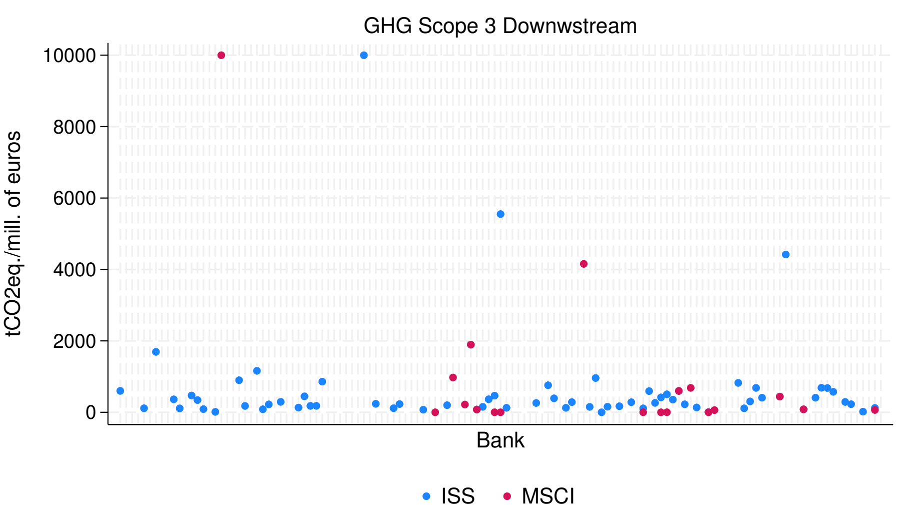
Papadopoulos, G. (2022), Discrepancies in corporate ghg emissions data and their impact on firm performance assessment, Technical Report 2022/12, JRC Working Papers in Economics and Finance, European Commission, Ispra.

Talbot, D. & Boiral, O. (2018), ‘Ghg reporting and impression management: An assessment of sustainability reports from the energy sector’, *Journal of Business Ethics* **147**.

Ángel Iván Moreno & Caminero, T. (2023), Assessing the data challenges of climate-related disclosures in european banks. A text mining study, Working Papers 2326, Banco de España.
URL: *https://ideas.repec.org/p/bde/wpaper/2326.html*

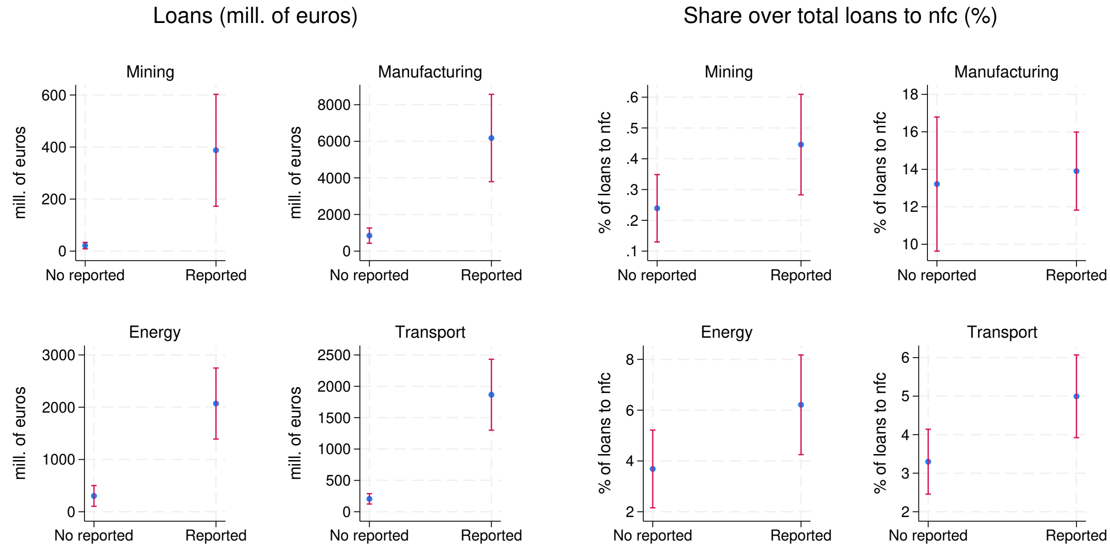
Appendix 1: Additional information on scope 3 emissions

Figure A1: Heterogeneity of scope 3 emissions footprint across data sources



Note: Data on banks' GHG emissions by different providers as of 2022. Each dot refers to a bank in the sample. The figure was set with an upward bound for the emissions data for graphic purposes. On the x-axis, banks are sorted according to their loan size to non financial corporations, so that the first bank on the left grants the smallest amount of loans, while the last one has the largest amount.

Figure A2: Self-reported scope 3 emissions availability and exposure to high-emitting sectors



Note: The figure displays the average (blue dot) amount of loans granted by the banks within our sample to a few high-emitting sectors (in millions of loans panel) and the average share of loans granted to such sectors over the bank's total loans to non-financial corporations in 2022 (in percentage) in 2022. Confidence intervals are in red. Two groups are considered: 1) the banks for which we have reported scope 3 emissions according to Bloomberg (Reported); 2) those for which we do not have this piece of information (No reported).

Table A1: Differences in banks' scopes 3 emission footprints across sources and exposure to high-emitting sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Mining	-91.91*	-118.8	-123.2	-1066.1	-127.0	-42720.3	-27892.9	-1044417.6	-27464.8	395.7***	-1486664.6	262.0	-1135220.8	-0.198
	(0.052)	(0.118)	(0.281)	(0.336)	(0.267)	(0.318)	(0.337)	(0.321)	(0.331)	(0.003)	(0.313)	(0.389)	(0.304)	(0.731)
Manufacturing	-1.796	-1.507	-8.788	-106.7	-9.596	-5722.5	-5902.5	-15997.1	-5594.8	-42.87*	-5101.4	-181.1	-52215.1	0.374
	(0.786)	(0.837)	(0.290)	(0.274)	(0.251)	(0.265)	(0.249)	(0.713)	(0.253)	(0.090)	(0.923)	(0.189)	(0.408)	(0.272)
Energy	8.239	-17.87	-26.35	-141.0	-19.96	-24070.0	-6121.4	-350797.6	-5935.3	-53.88	-548802.9	-213.1	-580540.5	0.965
	(0.770)	(0.472)	(0.521)	(0.679)	(0.643)	(0.232)	(0.349)	(0.334)	(0.342)	(0.274)	(0.330)	(0.207)	(0.310)	(0.454)
Transport	-5.295	8.375	11.11	-87.92	9.165	12069.3*	9943.4	255619.5	9775.2	215.7**	308590.8	148.5	351007.3	-0.653
	(0.759)	(0.517)	(0.450)	(0.151)	(0.556)	(0.072)	(0.128)	(0.352)	(0.131)	(0.025)	(0.347)	(0.192)	(0.320)	(0.362)
Observations	156	149	170	70	170	181	203	76	209	221	74	232	113	351
R^2	0.105	0.066	0.067	0.106	0.078	0.090	0.057	0.088	0.055	0.444	0.111	0.099	0.125	0.020

Note: Results from a set of multivariate regressions where the dependent variable is the difference in the bank's scope 3 emissions footprints between two different sources. The independent variables refer to the share of bank's loans to non-financial corporations belonging to high-emitting sectors (mining, manufacturing, energy and transport). Each column refer to a different pair of data sources. All the regressions include the constant and a year fixed effect. Errors are clustered at the bank level. P-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix 2: Analysis of GenAI-based emission data

This Appendix discusses in greater details the replies provided by Claude to our prompt. First, Claude 3.5 provided us with a table displaying the relevant data (hereafter, 1st run) by clarifying that the supplied data was not extracted from the banks' sustainability reports, but it was estimated to be representative of typical emissions for these institutions and should be verified against the banks' most recent official reports.[15] The values were not changed once we specified that, as investors, we sought to assess the carbon footprint of the banks in the list. However, the tool further stressed that the data consist of rough estimates that should be considered order-of-magnitude approximations at best and they are not suitable for investment decisions without verification against official sources.[16] Once Claude 3.5 was asked to assess the degree of confidence for the reported estimates in a scale from 1 (lowest confidence) to 10 (highest confidence), it assigned a higher confidence score to the data related to large global banks (i.e., between 6 and 7) since these banks have comprehensive public disclosures and more standardized reporting, making estimates more reliable; slightly lower scores for mid-sized European banks (i.e. 5 or 6) as they generally have good disclosure although differences in methodology and completeness; low values for smaller regional banks (between 3 and 4) with limited public data, more reporting inconsistencies and less standardized methodologies; minimum scores (2 or 3) for specialized institutions with very diverse business models that make comparison and estimation difficult, as well as for counties with less mature sustainability reporting frameworks. The tool further clarified that the confidence is particularly low for scope 3 emissions, which constitute the vast majority of a bank's carbon footprint but are reported using widely varied methodologies. Finally, when inquire, Claude 3.5 specified that it is unable to accurately perform the task of providing actual verified emissions data.[17]

When launching the prompt with Claude 7 (hereafter, 2nd run), instead, we specified to assume that we were investors or supervisory authorities trying to assess the carbon footprint of the banks in the list by using estimated or reported data of emissions. Claude 7 replied to our request by providing a table that displayed: 1) reported data based on publicly available sustainability reports and TCFD disclosures as of 2022, when available, and 2) estimated data, mostly for the smaller banks. For the scope 1 and 2 emissions, the estimates were based on the bank size (e.g., number of employees, total assets), location (e.g., country's energy mix, and data available for peer financial institutions). Scope 3 emissions, instead, were estimated using the credit and investment portfolios when available, using sectoral emission factors, and standard methodologies (i.e., the PCAF - Partnership for Carbon Accounting Financials), reported data by banks with similar size and business model or sectoral data for the banking sector. For scope 1 and 2 emissions, more than 57% of the data were classified as reported, while this was the case only for 17% of the scope 3 emissions. The reliability of the estimated scope 3 emissions was classified most of the times as medium (60% of the banks), while in almost all of the other cases the reliability score was low, especially for smaller banks with less public information for which the estimates were mostly based on sectoral averages.[18] Once again, the tool specified that these were not official data and need to be verified against official sources, as described above for Claude 3.5 (1st run).[19]

[15] It first included only a subset of the largest banks in the sample and then, when prompted, added the remaining ones.

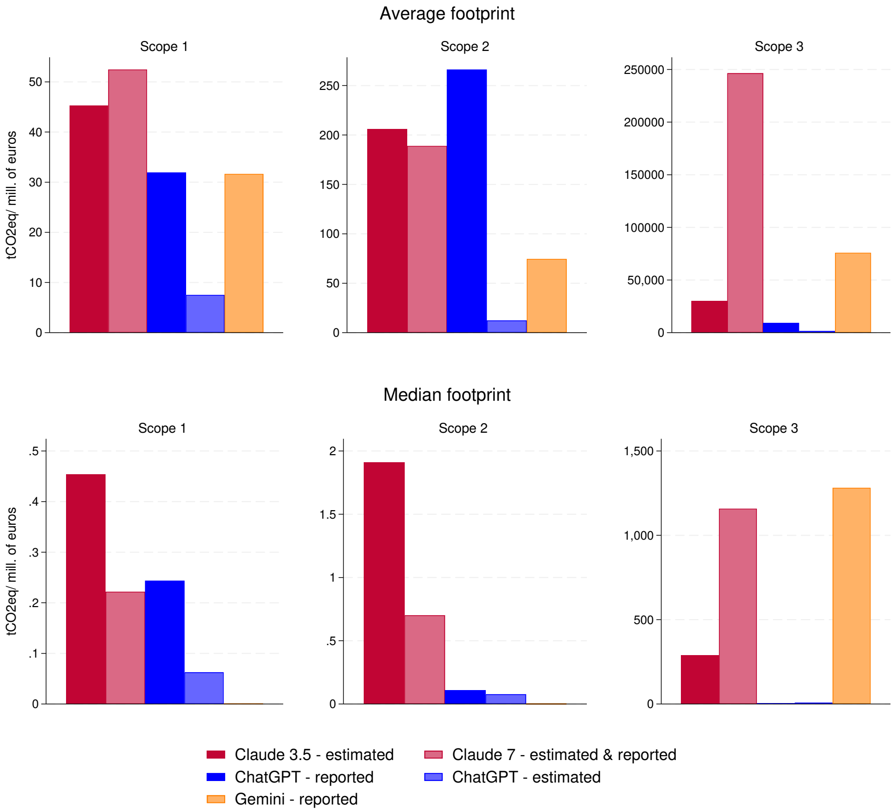
[16] It recommended that investors take a more systematic approach to obtain accurate emissions data. The recommendation included replacing estimated values with official and verified data from sustainability reports or other sources (i.e. CDP), assessing the precision of scope 3 emissions (particularly category 15, i.e., financed emissions), as these typically represent more than 95% of the total carbon footprint of banks and are most material for investors, and documenting the methodologies used by each bank.

[17] The following reasons were mentioned: 1) access limitations: it does not have direct access to subscription databases like CDP, Bloomberg, or MSCI ESG; 2) update of data: training data has limited information on 2022 and 2023 sustainability reports for many of these banks, which are necessary for investment decisions; 3) verification requirements: any data used for investment decisions should be carefully verified against official sources.

[18] The reliability score was medium for most of the banks whose available information on the portfolios was limited, but data for similar institutions allow to apply more detailed models.

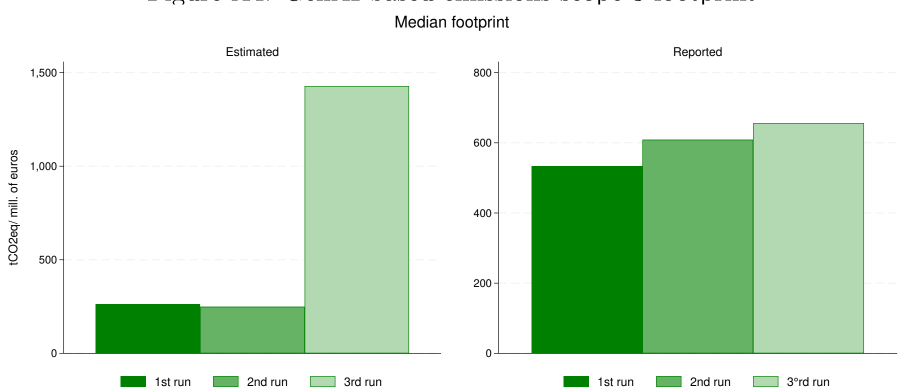
[19] Scope 1 and 2 emissions in the 2nd and 3rd runs are the same, as in the last run only the scope 3 emissions were

Figure A3: GenAI-based emissions footprint data over different runs



Note: For each bank, emission footprint is computed as the ratio between the bank's GenAI-emissions and the total loans to non-financial corporations as of the end of 2022.

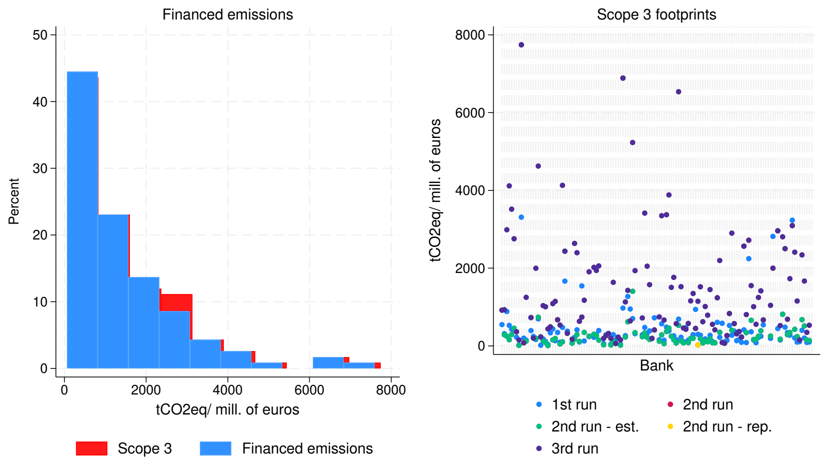
Figure A4: GenAI-based emissions scope 3 footprint



Note: Median scope emissions based on GenAI data for the sample of banks for which reported and estimated data are available in the 2nd run.

corrected.

Figure A5: GenAI-based financed emissions and bank-level heterogeneity across runs



Note: Data on banks' GenAi Scope 3 emissions as of 2022. The figures were set with an upward bound for the emissions data for graphic purposes. In the right-hand side plot, each dot refers to a bank in the sample. On the x-axis, banks are sorted according to their loan size to non-financial corporations, so that the first bank on the left grants the smallest amount of loans. In the left-hand side plot, the figure displays, the histograms of the total and financed scope 3 emissions in the 3rd run.

Table A2: GenAI scope 3 emissions and footprint, banks size and credit portfolio

	(1)	(2)	(3)	(4)	(5)
			Scope 3 emissions		
Tot. loans	946.1**	1039.7***	102.4**	1424.7**	1015.4**
	(0.025)	(0.009)	(0.016)	(0.027)	(0.011)
R^2	0.394	0.428	0.043	0.426	0.419
Obs.	129	129	107	22	129
			Scope 3 emissions footprint		
Tot. loans	-0.357	-0.584	-1.543	0.00204	-2.980
	(0.314)	(0.254)	(0.276)	(0.516)	(0.310)
R^2	0.003	0.004	0.005	0.031	0.003
Obs.	129	129	107	22	129
			Scope 3 emissions footprint		
Mining	-18058.5	-29246.5	-65761.0	237.6	-150279.5
	(0.315)	(0.258)	(0.279)	(0.174)	(0.312)
Manufacturing	-3399.9	-5464.6	-5624.9	-19.75	-28093.7
	(0.306)	(0.250)	(0.253)	(0.384)	(0.306)
Energy	-2869.5	-4596.5	-4554.8	7.106	-23683.6
	(0.311)	(0.258)	(0.258)	(0.898)	(0.311)
Transport	-3966.6	-6545.5	-6565.5	-157.9	-33081.9
	(0.327)	(0.264)	(0.265)	(0.105)	(0.323)
R^2	0.024	0.030	0.031	0.150	0.024
Obs.	129	129	107	22	129

Note: Results from a set of regressions where the dependent variables bank-level GenAI-based emissions data as of 2022 under the different runs (upper panel) or the scope 3 emissions footprint computed as the ratio between scope 3 emissions and the banks' total loans to non-financial corporations (middle and lower panels). *Tot. Loans* is the banks' total loans to non-financial corporations at a given year. In the lower panel, the independent variables are the share of bank loans to non-financial corporations belonging to high-emitting sectors (mining, manufacturing, energy and transport). All the regressions include the constant. Each column refers to a data source: (1) is the 1st run, (2) 2nd run, (3) 2nd run - estimated data, (4) 2nd run - reported data, (5) 3rd run. Robust standard errors. P-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$