Journal of Industrial Ecology (2026) 30:23–38
https://doi.org/10.1007/s44498-026-00003-5

RESEARCH

Completing the SME carbon profile: scalable prediction of scope 3 emissions

Alec Phillpotts[1,2] · Anne Owen[1] · Jonathan Norman[1] · Anna Trendl[2,3] · John Gathergood[4] · Norbert Jobst[5] · David Leake[2]

Received: 2 April 2025 / Accepted: 5 January 2026 / Published online: 20 March 2026

Abstract
Despite growing recognition of significance, a business's Scope 3 emissions remain rarely measured and as a result are poorly understood. This situation is particularly common amongst small and medium-sized enterprises (SMEs), which face additional obstacles to emission measurement. With this paper, we present a transaction-based approach to facilitate SME Scope 3 engagement. Using financial transaction data for 150,000 + UK SMEs, we produce spend-based Scope 3 estimates across key Greenhouse Gas Protocol categories. We then fit a series of hierarchical regression models to both quantify and identify firm-level Scope 3 emissions, with minimal user inputs. We find that this approach is effective in predicting the upstream emissions of both purchased goods & services (RSQ = 0.87) and fuel and energy-related activities (RSQ = 0.89 and 0.72), whilst weaker for more targeted categories such as business travel. We also find a small number of recurrent industry hotspots tend to account for 75% of a firm's upstream emissions. By leveraging objective, standardised data to estimate emissions, this method provides a low-input alternative to costly micro-studies for generating Scope 3 insights, extending the visibility of emissions beyond a firm's direct operations, revealing emission hotspots and supporting the development of value chain decarbonisation strategies.

Keywords Industrial ecology · Scope 3 · Emission modelling · Embodied emissions · Financial transaction data

1 Introduction

Business emission reporting has traditionally focused on the emissions produced directly by operations (Scope 1) and energy-use (Scope 2). Whilst these emissions are both relatively easy to quantify and directly under a business's organisational control, this neglects the substantial share of emissions that occur across the broader value chain (Scope 3) (Serafeim & Velez Caicedo, 2022). Typically these indirect emissions far exceed a firms combined Scope 1 and 2 totals (Huang et al., 2009). Overlooking Scope 3 emissions both limits business decarbonization efforts (Hettler & Graf-Vlachy, 2024) and risks inefficient policy interventions (Hertwich & Wood, 2018).

Recognition of the importance of Scope 3 emissions is gaining momentum. Policy developments to integrate Scope 3 into disclosure frameworks are underway in Europe (DESNZ, 2023; EU, 2022), while Australia is phasing in the introduction of mandatory Scope 3 emissions reporting for larger businesses. However, extending these frameworks to small and medium-sized enterprises (SMEs) requires careful consideration.

Scope 3 emissions are complex relative to Scope 1 and 2 emissions (Cheema-Fox et al., 2021). Their calculation necessitates subjective decisions on system boundaries, whilst requiring extensive external data on supplier locations, production techniques, and knowledge on various other factors within global supply chains (Acquaye et al., 2014). Under the Greenhouse Gas (GHG) Protocol, Scope 3 emissions are divided into 15 distinct and mutually exclusive

✉ Alec Phillpotts
ee21ap@leeds.ac.uk

1 Sustainability Research Institute, University of Leeds, Leeds, UK

2 Behavioural Science, Lloyds Banking Group, London, UK

3 Warwick Business School, University of Warwick, Coventry, UK

4 School of Economics, University of Nottingham, Nottingham, UK

5 Portfolio Analytics, Lloyds Banking Group, London, UK

categories, capturing embodied upstream emissions from purchased goods and services, as well as downstream emissions from sold goods and services, including distribution, use, and end-of-life treatment (WRI & WBCSD, 2020).

For larger firms with complex supply networks, significant regulatory exposure, and notable individual environmental impact, firm-specific Life Cycle Assessments (LCAs) are the preferred approach. LCAs involve bottom-up techniques to account for the specific processes, suppliers, and regions involved in each stage of the life cycle. The resulting data are primary, and therefore considered more reliable (Busch et al., 2022). Whilst guided by international standards, and adaptable to the GHG Protocol Scope 3 categories, LCAs are inherently resource-intensive and methodologically complex, even for dedicated teams within large corporations (Goldhammer et al., 2017). Calculating these emissions requires technical knowledge, and involves significant operational and transactional costs (Isil & Sebastianelli, 2020). Moreover, differences in methodological approaches undermine data comparability (Robinson et al., 2015), while frequent errors further compromise data quality (Patchell, 2018).

To mitigate the substantial operational and transactional costs of primary data collection, firms often turn to secondary data sources. These can include pre-existing LCAs, or industry-average emission conversion factors derived from environmentally extended multi-regional input–output (EEMRIO) models (Nguyen et al., 2023). Whilst simpler, these approaches still require a level of technical expertise, and for some, the reliance on secondary factors introduces inaccuracies (Busch et al., 2022).

Despite this, different levels of accuracy and completeness are necessary to address different needs (Huang et al., 2009), with process-based data alone insufficient for producing robust and comparable estimates at scale (Minx et al., 2008). For SMEs, sole reliance on expensive micro-analysis presents a substantial barrier to the widespread adoption of Scope 3 emission measurement.

In our earlier work (Phillpotts et al., 2025), we build the argument for applying simple statistical models to proxy Scope 1 and 2 emissions amongst SMEs. Our approach uses large-scale financial transaction data (FTD) to fit models that are simple to apply, requiring two inputs: turnover and industry. This efficient, and highly scalable approach addresses the data gap created by the exclusion of SMEs from current reporting requirements. FTD-based emission estimates capture a simplified version of emissions, and its validation against actual emissions is challenging due to the scarcity of alternative data. However, we argue that transparent, standardised, and objective measurement represents a justified trade-off, at the expense of some absolute accuracy.

This paper builds on these foundations by extending the approach to SME Scope 3 emissions. Existing approaches to model unreported Scope 3 emissions have used regression and machine learning techniques (Carbon Disclosure Project, 2020; Nguyen et al., 2023; Serafeim & Velez Caicedo, 2022; Shakdwipee & Lee, 2016). These models rely on self-reported data for their training, presenting inherent limitations; inconsistent methodologies across reporting firms undermine comparability, while the training data underrepresents SMEs. With this paper, we utilise the FTD of over 150,000 UK SMEs to estimate Scope 3 emission categories, creating a large dataset for a previously unobserved segment of the business population.

Drawing on our novel, micro-level dataset, we make two contributions: first we explore how different transaction categories contribute to a firm's Scope 3 emissions, and second, we develop a series of statistical models to quantify and identify Scope 3 emissions. Through this work, we demonstrate how FTD can support the mainstreaming of Scope 3 emission measurement and introduce a simple benchmarking tool to generate Scope 3 emissions insights. Thereby transforming extensive, granular, otherwise private data into accessible macro-level insights to facilitate a wider adoption of Scope 3 measurement amongst smaller, often under-engaged actors (British Business Bank (BBB), 2024).

The remainder of this article is structured as follows. First, we describe the data and method used to calculate Scope 3 categories, identify hotspot thresholds, and outline the statistical techniques applied in developing our models. Next, we present model results, interpretations, and stability tests. Finally, we conclude with a discussion of the key findings, their practical implications, limitations, and recommendations for future research.

2 Materials and methods

In partnership with Lloyds Banking Group (LBG), one of the UK's largest retail banks (LBG, 2023), this study accessed a substantial dataset of transactions made by UK SMEs in the calendar year 2021. Here, we define SMEs as firms with an annual turnover of less than £36 million (Companies Act, 2006).

Our methodological approach is outlined in the following subsections, with Fig. 1 providing a schematic overview of the Scope 3 extension to the model. First, FTD is obtained from LBG, and an iterative sample restriction process is implemented. The FTD from this sample is then used to estimate Scope 3 emission categories. We then produce prediction models by fitting a series of hierarchical regression models to the FTD-based emission estimates, using firm-level predictors from LBG, and industry predictors from public accounts when required. In some instances, the

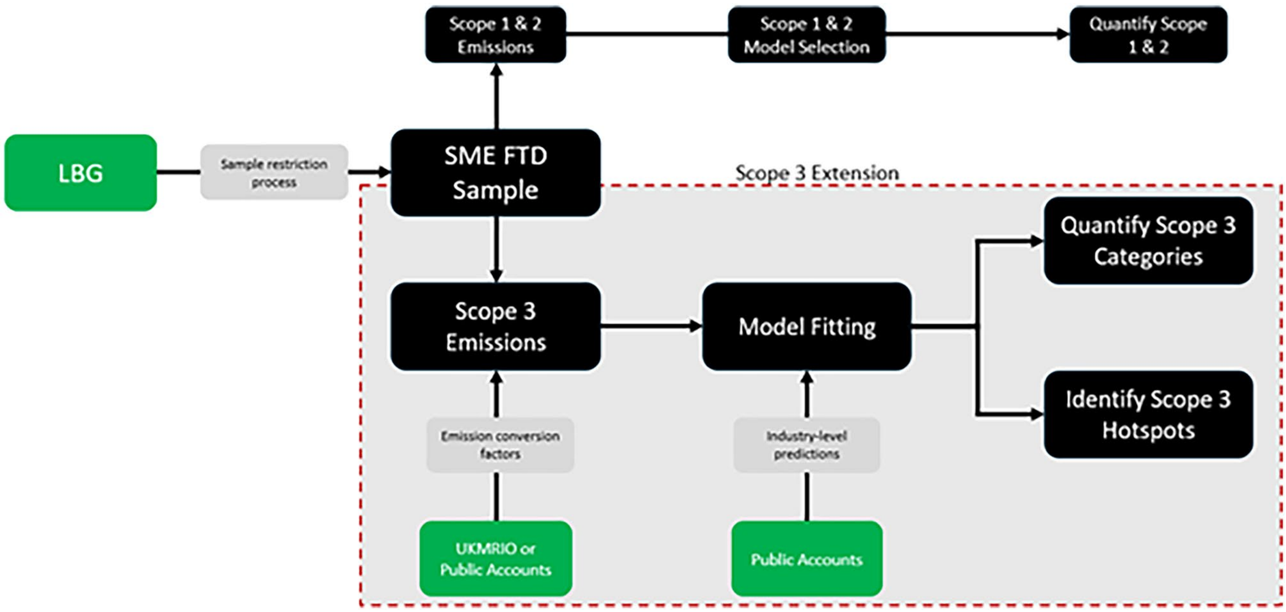

Fig. 1 Overview of prediction model development. Black rectangles indicate key stages of the modelling process, green denotes data sources, and grey highlights the specific inputs and processes required at each step. Adapted from Phillpotts et al. (2025).

model fitting process requires additional sample restrictions to ensure model convergence. For each category, we evaluate model performance and select a preferred specification, before conducting further validation tests.

2.1 Sample selection

A series of sample restrictions are required to ensure our analysis is based on comprehensive financial data histories, isolating SMEs who use LBG as their primary banking provider, and excluding those with significant financial activities through other institutions (Phillpotts et al., 2025). A full account of this process, including rationale and impact of each step is provided within Supplementary Information SI1.

We also restrict our sample to only SMEs for which FTD can reasonably serve as a proxy for both firm revenue and Scope 3 emission-generating activities. We therefore exclude SMEs in sectors including agriculture, energy, and water, where substantial industry-specific Scope 3 emissions require more tailored modelling approaches. Certain service-sector SMEs are also omitted, such as those in insurance, real estate, and legal or accounting services, due to frequent handling of non-turnover credit lines causing a systematic overestimation of turnover (Phillpotts et al., 2025). Table S2 in Supplementary Information SI1 provides a full list of explanations of exclusions where relevant, along with the sample and population figures for each industry.

With these restriction steps in place, our sample consists of 167,169 companies across 57 industries. Collectively, these companies account for an estimated £103 billion in annual revenue, with over 101 million financial transactions collectively in 2021.

2.2 FTD-based emissions estimates

2.2.1 Mapping to greenhouse gas (GHG) protocol categories

The GHG Protocol (WRI and WBCSD, 2020) defines 15 categories of Scope 3 emissions. These categories are intended to encompass all indirect emissions from business activities. Consequently, a firm's FTD does not directly translate to all Scope 3 categories, and assumptions may be required to enable compatibility. Table 1 introduces each Scope 3 category, noting its coverage by FTD and the need for any assumptions in its estimation.

2.2.2 Scope 3 emission conversion factors

The primary source for Scope 3 emission conversion factors is the UKMRIO (Owen & Kilian, 2024). The UKMRIO is a Single-country National Accounts Consistent (SNAC) EEMRIO model, used to produces the official annual consumption-based emissions for the UK, making it ideal for a UK focused methodology (Tukker et al., 2018). EEMRIO conversion factors reflect the embodied environmental load of a unit of final demand (e.g. kg CO_2e / £) (Wiedmann et al., 2011), a full derivation of which, can be found in Kitzes (2013, pp. 497–498). Whilst these conversion factors

Table 1 GHG Protocol Scope 3 categories and their alignment with FTD

Scope 3 Category	Captured by FTD?		Included in Model?	
1. Purchased goods & services	Yes	FTD reflects the goods and services a business purchases when paid via the business account. These transactions can be matched to emissions using appropriate emission factors	Yes	
2. Capital goods	Partially	Capital asset purchases are captured in FTD; however, without additional context, FTD alone cannot reliably distinguish capital goods from regular purchases, nor identify the specific type of capital asset involved	No	
3. Fuel- and energy-related	Yes	Fuel and energy purchases are visible in FTD and can be used to calculate Scope 1, 2 and associated Scope 3 emissions	Yes	
4. Upstream transport & distribution	Partially	Payments to logistics providers are visible when made directly. However, transport and delivery fees are often bundled within purchases and may not be separately identifiable	Yes	We assume upstream transport and distribution costs are included within the purchases of goods. These emissions are therefore included within Category 1
5, Waste from operations	Partially	Payments to waste service providers are visible. However, FTD does not capture waste volume, type, or treatment method. Additional data or assumptions are beneficial for measurement	No	
6. Business travel	Partially	Travel-related purchases such as flights, hotels, and car hire are identifiable in company FTD. However, expenses that are paid by employees and later invoiced to the company are not captured and thus remain unmeasured	Yes	We assume all business travel spend in our sample is made through the business account
7. Employee commuting	No	Commute-related spending is made by employees and does not typically appear in business transactions	No	
8. Upstream leased assets	Partially	Lease payments are visible, but emissions depend on asset type and use, which cannot be determined by FTD alone	No	
9. Downstream transportation and distribution	Partially	Courier and distribution payments are visible when made directly. However, as with upstream transport, bundled costs and lack of directionality (up vs. downstream) can make attribution difficult	Yes	We assume all payments to courier, distribution and logistics firms are related to the downstream transportation of goods and services
10. Processing of sold products	No	No direct purchase associated	No	
11. Use of sold products	No	No direct purchase associated	No	
12. End-of-life treatment of sold products	No	No direct purchase associated	No	
13. Downstream leased assets	No	Lease income may appear in FTD, but usage-based emissions are not measurable with FTD alone	No	
14. Franchises	No	Payments such as franchise fees or royalties may be visible, but operational emissions would be challenging to estimate. Moreover, SMEs are unlikely to have associated franchises, due to their size	No	
15. Investments	No	Investment income or flows may appear, but emissions from investments require detailed portfolio data. In addition, SMEs, are unlikely to be holding large sums of financial investments	No	

are generally attributed to a unit of final demand, the Leontief inverse assumes sectoral goods are homogenous and thus have the same cradle-to-gate emissions per monetary unit. As such, intermediate inputs have the same unit footprint as products sold to the final consumer (Hertwich & Wood, 2018). The underlying assumption of EEMRIO conversion factor use is that the supply structure of each purchase can be approximated by the corresponding economic sector, as represented in the UK Supply and Use Tables (ONS, 2022; Schmidt et al., 2022).

The UKMRIO produces consumption-based conversion factors for 112 industries, denoted by their Standard Industrial Classification of economic activities (SIC code). To make these conversion factors compatible with FTD, SIC industries are mapped to all but two headings of spend: “Energy & Utilities” and “Vehicle Fuelling”. The mapping table for this process can be found in Supplementary Information SI2.

Spend categorised under “Energy & Utilities” and “Vehicle Fuelling” can instead be combined with direct conversion factors to estimate Scope 1 and Scope 2 emissions (Phillpotts et al., 2025) or as in this case, well-to-tank (WTT) conversion factors. The resulting estimates reflect the emissions produced during the extraction, processing, and distribution of fuels and energy. Table 2 displays the resulting interquartile range and median emissions per estimated Scope 3 category, along with observation numbers, whilst further details of the conversion factor calculations and their mapping to transactions are provided in Supplementary Information SI3.

2.2.3 Defining scope 3 emission hotspots

While some Scope 3 categories are narrow, Category 1. Purchased Goods and Services span diverse, sector-specific emission sources. To compare patterns across firms and sectors, we express each emission source as a proportion of total calculated Scope 3 emissions and use these proportions to identify recurrent hotspots. Table 3 displays different minimum-contribution thresholds, showing the number of expenditure codes used at the industry level, the number used per firm, and the share of emissions captured. For each metric we report the mean, median, and standard deviation.

Using a 10% contribution threshold, we capture, on average, 74.5% of emissions while reducing the median number categories from 24 to 2 per firm. We define hotspots as categories contributing above this 10% threshold and visualise industry mean values in Fig. 2.

2.3 Hierarchical regression models

To develop models for predicting Scope 3 emissions, we fit a series of hierarchical regression models to the FTD-based emissions estimates. To maintain the statistical reasoning outlined in (Phillpotts et al., 2025), turnover representativeness and adequate group sizes are ensured (Ali et al., 2019).

2.3.1 Quantifying scope 3 emission categories

To quantify estimates for Scope 3 emission categories, we utilise a hierarchical linear regression specification, allowing for industry-specific effects to vary (Gelman & Hill, 2006). Following Goldhammer et al. (2017), we specify a log–log relationship, allowing for coefficients to be interpreted as elasticities.

2.3.1.1 Variable selection To predict Categories 1, 6 and 9, we rely on a selection of firm-level variables, outlined in Table 4. Additional variables to capture Scope 3 emissions would need to reflect more granular operational details, such as supplier characteristics or transportation metrics on distance or freight mode (Wang and Ye, 2025). Standardised versions of these inputs are rarely available for SMEs at scale, and the inclusion of weak proxies risk introducing measurement error and biasing coefficient estimates (Hausman, 2001).

For Category 1, we retain the full sample, whilst Categories 6 and 9 require further targeted sample refinement to support model convergence. Here, we exclude SMEs with zero expenditure in the relevant category (43% of the sample for downstream transportation, and 17% for business travel) and remove extreme values by trimming the top and bottom 10% of category spend, relative to total expenditure. The resulting samples comprise 111,098 SMEs for Category 6 and 75,622 for Category 9, with industry coverage maintained at 57 industries.

As Category 3. Fuel- and energy-related activities emissions should directly align with a firm’s Scope 1 and 2 emissions, the model specifications developed and validated in Phillpotts et al. (2025) for Scope 1 and 2 emissions are applied, adapting each to estimate WTT emissions of fuels and electricity respectively. Here, models rely on turnover and 2-digit SIC code, as well as variables on industry-level Scope 1 and 2 emission intensity, as well as the industry distribution of turnover.

2.3.1.2 Model fitting and evaluation Models for Categories 1, 6 and 9 are built sequentially, starting with two and increasing to five variables, evaluating the required complexity to explain the dependent variables (Cohen et al., 2013). Under a transaction-based methodology, higher Scope 3 emissions

Table 2 Number of observations, interquartile and median emissions by aggregated industry and Scope 3 category. An additional table at a disaggregated industry level can be found in Supplementary Information SI3

SIC Section	Sample	Category 1. Purchased Goods and Services			Category 3.1 Fuel-related activities			Category 3.2 Energy-related activities			Category 6. Business Travel			Category 9. Downstream transportation and distribution		
		Q1	Median	Q2	Q1	Median	Q2	Q1	Median	Q2	Q1	Median	Q2	Q1	Median	Q2
	#	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*	*kg Co2e*
A - Agriculture	696	2,139	57,261	160,016	1,872	3,650	10,601	207	429	973	577	1,286	2,689	45	308	1,821
C- Mining	78	61,734	142,245	460,682	2,637	6,951	23,407	232	647	2,391	771	1,511	3,557	423	3,815	18,208
F - Manufacturing	15,497	20,591	60,863	197,533	1,296	2,733	8,350	65	203	597	416	974	2,115	89	827	5,701
G - Construction	44,901	14,158	31,430	88,309	2,153	4,358	9,775	199	427	1,012	500	1,031	1,990	14	83	805
H - Wholesale & Retail	34,107	21,483	63,890	196,064	1,736	5,187	16,430	293	690	1,810	228	602	1,419	70	489	3,831
I - Transportation	2,244	2,883	8,100	30,409	3,261	5,991	10,280	693	1,376	2,417	495	1,260	2,942	6,244	22,753	83,916
J - Hospitality	16,897	17,642	43,888	92,906	403	810	1,748	163	405	991	202	683	3,320	10	54	369
M - Prof Activities	3,317	9,094	20,636	67,617	909	1,733	3,790	164	381	891	274	656	1,545	17	79	593
N - Admin Services	7,127	8,682	19,281	60,308	1,104	2,253	4,847	98	229	502	340	787	1,644	15	54	377
P - Public	9,171	4,792	11,540	34,089	2,510	18,210	78,787	567	2,712	9,364	350	794	1,620	12	56	461
Q - Education	5,419	6,604	17,150	53,743	2,024	4,120	9,067	324	716	1,706	774	1,368	2,666	14	81	905
R - Health	11,953	8,996	38,942	136,825	1,176	2,529	7,312	222	501	1,580	191	501	1,212	16	52	194
S - Arts	6,388	11,397	41,184	131,622	700	1,086	1,721	181	280	448	142	369	871	45	56	249

Table 3 The impact of minimum contribution thresholds on Scope 3 emission sources

Min Scope 3 Contribution	Industry Expenditure Codes			Firm Expenditure Codes			Emission Captured		
	Mean	Median	Std	Mean	Median	Std	Mean	Median	Std
	#	#	#	#	#	#	%	%	%
None	72.3	74	15.6	26.8	24	13.0	100%	100%	0
1%	21	21	3.2	8.8	8	3.5	95.7%	96.1%	0
5%	8.8	9	1.1	3.6	3	1.5	83.7%	84.9%	0.1
10%	**5.6**	**6**	**0.8**	**2.3**	**2**	**0.9**	**74.5%**	**76.2%**	**0.1**
20%	3.9	4	0.5	1.4	1	0.6	61.7%	64.6%	0.2

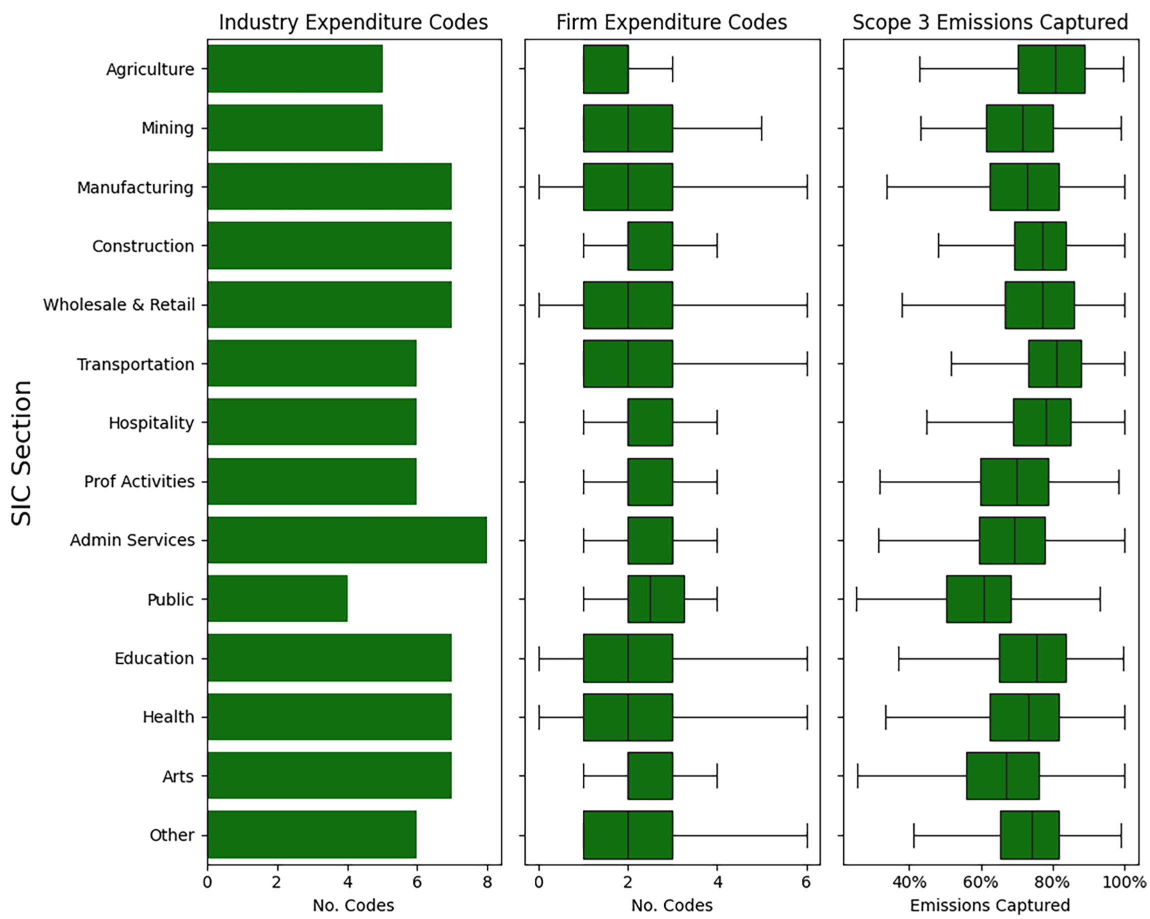

Fig. 2 Industry breakdown of hotspot spend categories and the level of emissions captured by them. Underlying data for this figure can be found in Table S1 of Supplementary Information SI7

are associated with both greater absolute levels of expenditure and a greater allocation of spend to high intensity categories. We first reflect these dynamics with just turnover and industry classification in Model 1, allowing the effect of turnover to vary by industry. We then extend the model by nesting additional three-digit SIC codes within the industry groupings (Model 2) and subsequently incorporate capital variables to capture further firm-level heterogeneity (Model 3). These variables are specified as fixed effects to mitigate convergence issues. The hierarchical regression structure of these models are as follows:

$$Model\,1 : \ln y_i = \left(\beta_t + \beta_{t,j(i)}\right) \ln turnover_i + \varepsilon_i$$

$$Model\,2 : \ln y_i = \left(\beta_t + \beta_{t,j(i)} + \beta_{t,k(i)}\right) \ln turnover_i + \varepsilon_i$$

Table 4 Predictor variables for the hierarchical regression model. Adapted from Phillpotts et al. (2025)

Variable	Unit	Description	Calculation
Turnover	£	Turnover serves as an indicator of a firm's operational scale and is included in the analysis because, all else being equal, higher turnover typically correlates with greater emissions	Credit turnover is derived by summing the credit transactions recorded within a firm's bank account in the year (excluding identified non-turnover credit lines)
2 Digit SIC Code	*SIC*	The Standard Industrial Classification (SIC) identifies a firm's primary business activities. Including this industry identifier allows for more accurate comparisons of expected emission intensity among firms performing similar activities	SIC classifications are the standard industry categories in the UK with official economic and environmental data commonly published on a 2 digit SIC code basis, whilst it is standard practice for firms banking with LBG to be assigned SIC codes at both the 2- and 3-digit level by relationship managers.
3 Digit SIC Code	*SIC*		
Capital Spend – Assets	%	Capital expenditure reflects a firm's investment in assets. This variable is included because high capital spending on newer energy-efficient technologies may serve to lower a firm's emission intensity. For ease of calculation, we define capital expenditure as purchases that require financing	Asset capital spend is calculated by summing the percentage of total spend, spent on asset financing and plant hire for the three-year period ending 2021. We take an average of these percentages to gain a single figure, capturing asset spending across the three-year window
Capital Spend – Vehicles	%		Vehicle capital spend is calculated by summing the percentage of total spend, spent vehicle financing for the three-year period ending 2021. As above, we calculate a single figure, capturing vehicle spending across the three-year window

$$
\begin{aligned}
Model\,3:\ \ln y_i =& \beta_t \ln turnover_i + \beta_2 \ln assets_i \\
&+ \beta_3 \ln vehicles_i + \beta_{t,j(i)} \ln turnover_i \\
&+ \beta_{t,k(i)} \ln turnover_i + \varepsilon_i
\end{aligned}
$$

where:

y_i represents the absolute emissions each category of firm i's Scope 3 emissions.

ln represents the natural logarithm of the variable.

β_t represents the fixed coefficients for firm turnover.

$\beta_{t,j(i)}$ represents the variable coefficient for industry j of firm i.

$\beta_{t,k(i)}$ represents the variable coefficient for the additional industry grouping (k) of firm i.

β_2 represents the fixed coefficients for the firms' assets variable.

β_3 represents the fixed coefficients for the firms 'vehicles variable.

ε_i represents the residual error term.

Model specifications for estimating Category 3 are taken directly from Phillpotts et al. (2025), this emissions category is therefore split between Category 3.1 (fuels) and 3.2 (electricity), whilst only results for the final model is presented.

Fuels (WTT emissions related to Scope 1):

$$
\begin{aligned}
\ln y_i =& \beta_1 \ln intensity_i + \beta_2 \ln skew_i \\
&+ \beta_3 \left(\ln intensity_i \cdot \ln skew_i\right) \\
&+ \beta_{t,j(i)} \ln turnover_i + \varepsilon_i
\end{aligned}
$$

Electricity (WTT emissions related to Scope 2):

$$
\begin{aligned}
\ln y_i =& \beta_1 \ln intensity_i + \beta_2 \ln skew_i \\
&+ \beta_3 \left(\ln intensity_i \cdot \ln skew_i\right) \\
&+ \beta_{t,j(i)} \ln turnover_i + \varepsilon_i
\end{aligned}
$$

where:

y_i represents the absolute emissions of firm i's fuel or energy Scope 3 emissions.

ln represents the natural logarithm of the variable.

$\beta_{1,2,3}$ represents the fixed coefficients for industry variables.

$\beta_{t,j(i)}$ represents the variable coefficient of turnover for industry j of firm i.

ε_i epresents the residual error term.

For each model, we present coefficients and statistical significance, indicating the strength and direction of predictor relationships to the dependent variables, along with their standard errors, reflecting the estimate uncertainties. We evaluate both model fit (RSQ) and model fit with respect to complexity (AIC). We then select a preferred model and test out-of-sample performance using an 80/20 stratified train-test split with fivefold cross-validation (James et al., 2021).

2.3.2 Identifying scope 3 emission hotspots

Next, we fit a hierarchical binomial logistic regression framework on the proportional emissions data of Table and Fig. 2. This approach enables the estimation of an outcome probability (Gelman & Hill, 2006), thereby allowing us to identify and present hotspot categories based on firm-level predictors. Here, our dependent variable is defined as the log odds of an expenditure category contributing over 10% to calculated emissions. We run an iteration of this regression for each of the 624 expenditure categories.

The hierarchical binomial logistic regression analysis is computationally intensive due to the large sample size, industry groupings, and the 624 outcome iterations. Increasing model complexity posed challenges for model convergence and runtime efficiency. To address this, we implement two model formulas for each spend category. First, Model 4 predicts the log odds of hotspot occurrence given firm industry. Second, Model 5 predicts the log odds of hotspot occurrence given both firm industry and turnover, allowing for changes in hotspot likelihood with firm size.

As we run 624 iterations of each model, we report the average performance of each model specification. In addition to the AIC, we report the median residual as a measure of central tendency, where values close to zero indicate accurate predictions. We also present the model's log-likelihood, reflecting explanatory power, with higher values indicating a better fit. Finally, we convert the predicted log odds into probabilities, to enable visualisation and interpretation. Full model specifications, results and probability calculations can be found in Supplementary Information SI3.

3 Results

3.1 Hierarchical linear regression results

Summaries for the three models used to predict Categories 1, 6 and 9 are presented in Table 5. In Model 1, turnover returns a strong, positive coefficient. The elasticity is highest for Category 1 (0.87), whilst lower for Category 6 (0.55) and Category 9 (0.47). Explanatory power varies substantially, with RSQ values ranging from 0.87 for Category 1 to 0.55 and 0.41 for Categories 6 and 9, reflecting differences in how well turnover and industry alone explain each emissions source.

The introduction of more granular industry detail in Model 2 expands groupings from 58 to 212 SIC codes. These groupings slightly improve model fit across all categories, as seen in lower AIC values and small gains in RSQ.

The inclusion of asset and vehicle variables produces mixed effects across categories. Turnover coefficients remain largely unchanged across each category's models, with additional disaggregation and predictors unable to weaken its strong relationship with emissions. For Category 1, both additional variables have small negative coefficients, slightly reducing emissions after controlling for turnover. For Category 6, the vehicles variable shows a small positive relationship, while assets have no significant effect. For Category 9, the effect of the two additional variables is strongest, with elasticities of -0.07 and -0.12, suggesting that firms with greater in-house capacity may rely less on outsourced distribution services.

Models for both Categories 3.1 and 3.2 return identical RSQ values to what is observed in Phillpotts et al. (2025), indicating that changes in emission factors did not introduce unexpected variation in model performance. Across both models, all coefficients retain consistent correlations, with only slight changes in magnitude, reflecting the absolute differences between Scope 1 and 2 emissions and their WTT counterparts. The foremost difference lies in the skew term where the previously insignificant coefficient for electricity now returns as negative and highly significant (-0.54***).

The summaries in Table 5 indicate that the inclusion of additional firm-level variables did not significantly improve model performance, despite increasing model complexity and data requirements. Model 3 delivers the best or near-best fit across all categories but gains in explanatory power over earlier models are modest, just one to two percentage points higher than the baseline Model 1.

For Category 1, a simple and parsimonious model provides strong predictive power while remaining practical to implement, these emissions are well explained by a firm's turnover, reflecting scale, and its industry, capturing spending patterns. Model performance is further supported by the standardised calculation of emissions and broad emission category. In contrast, for Categories 6 and 9, explanatory power is weaker, with the most complex model failing to significantly improve performance. This reflects the higher uncertainties introduced by the necessary assumptions underlying emissions estimation.

Alongside the models for Categories 3.1 and 3.2, Model 1 is selected as the preferred specification for the remaining categories, and five-fold cross-validation is conducted to evaluate model robustness. The results indicate consistent performance across folds, with minimal variation in both RSQ and median errors. Category 1 shows the strongest performance, with a median absolute error of 8.5 kg CO_2e (28%), while Categories 3.1 and 3.2 produce smaller absolute errors (1.0 kg CO_2e and 0.2 kg CO_2e, respectively) reflecting the low magnitude of WTT emissions but higher relative percentage errors (both around 49%). Categories 6 and 9 exhibit weaker fits, with median errors of 0.4 kg CO_2e (59%) and 0.18 kg CO_2e (93%), respectively. Overall, the stable RSQ values across folds indicate that, regardless of absolute fit strength, the models capture relationships that

Table 5 Model summary. Asterisks indicate significance level where P values < 0.05 = *, < 0.01 = **, and < 0.001***

	Model 1		Model 2		Model 3	
	Coef	Std	Coef	Std	Coef	Std
Category 1. Purchased goods & services						
ln turnover	0.87***	0.01	0.87***	0.01	0.86***	0.01
ln assets					-0.02***	0.00
ln vehicles					-0.01***	0.00
RSQ (fixed)	0.74		0.75		0.75	
RSQ (total)	0.87		0.88		0.88	
AIC	269,421		248,957		247,866	
Category 3.1. Fuel-related activities						
ln turnover					0.77***	0.02
ln scp1_intensity					1.35***	0.21
ln skew					-0.82***	0.02
ln (intensity * skew)					-0.46***	0.05
RSQ (fixed)					0.41	
RSQ (total)					0.89	
AIC					80,304	
Category 3.2. Electricity-related activities						
ln turnover					0.78***	0.01
ln scp2_intensity					-1.07***	0.10
ln skew					-0.54***	0.03
ln (intensity * skew)					0.28***	0.03
RSQ (fixed)					0.59	
RSQ (total)					0.72	
AIC					210,345	
Category 6. Business Travel						
ln turnover	0.55***	0.01	0.54***	0.01	0.55***	0.01
ln assets					-0.00	0.00
ln vehicles					0.04***	0.00
RSQ (fixed)	0.27		0.28		0.27	
RSQ (total)	0.55		0.56		0.55	
AIC	321,419		318,082		317,793	
Category 9. Downstream transportation and distribution						
ln turnover	0.47***	0.02	0.47***	0.02	0.46***	0.02
ln assets					-0.07***	0.00
ln vehicles					-0.12***	0.01
RSQ (fixed)	0.07		0.07		0.09	
RSQ (total)	0.41		0.43		0.43	
AIC	322,529		321,169		320,393	

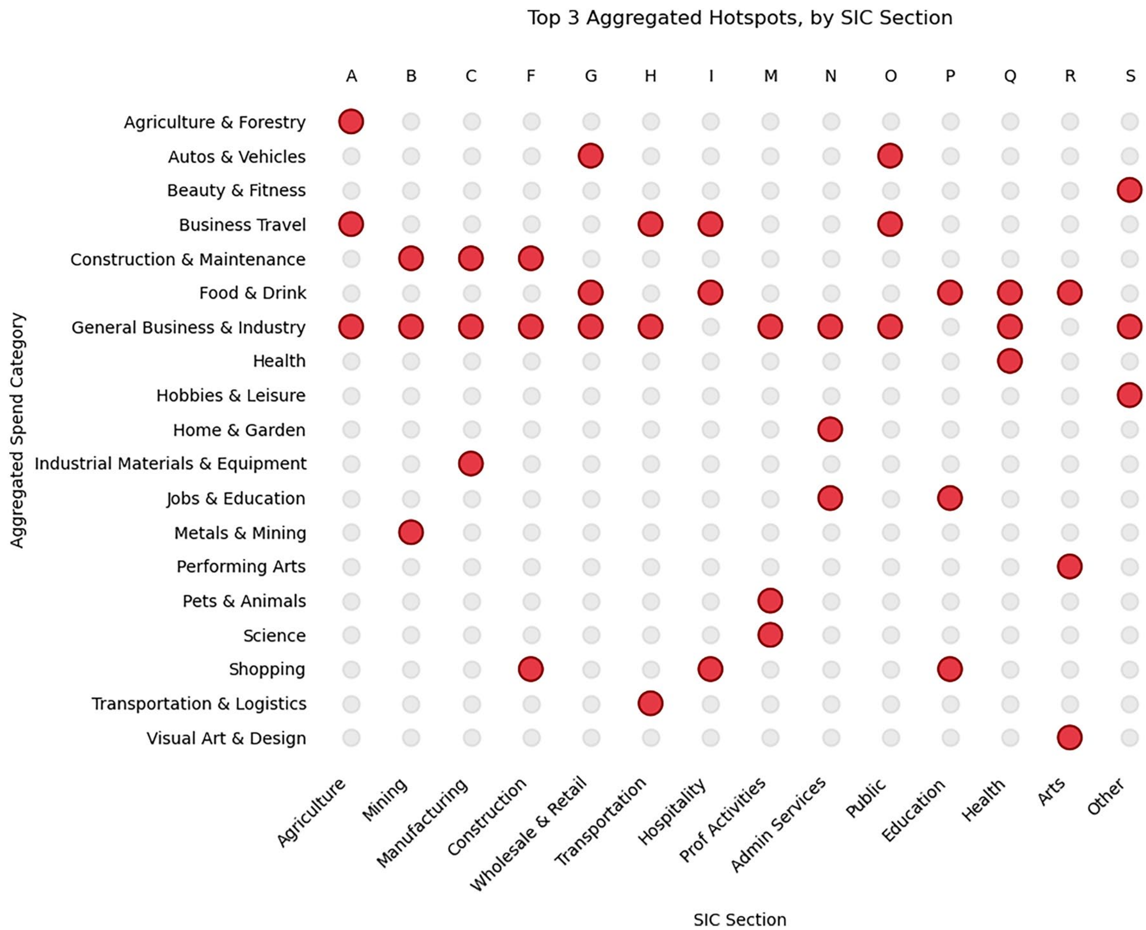

Fig. 3 Hotspot matrix showing the likeliest hotspot spend category, by SIC section. Underlying data for this figure can be found in Supplementary Information SI5

are reproducible and not dependent on specific training samples. Test and train results for each iteration can be found in Supplementary Information SI4.

3.2 Hierarchical binomial logistic results

Table S7 in Supplementary Information SI4 presents average summaries for the binomial regression iterations. Under Model 4, we observe an average median residual of −0.0018, while average log-likelihood is −2,562. The inclusion of turnover in Model 5 improves all performance metrics, reducing the median residual slightly to −0.0015 respectively, and increasing the average log-likelihood to −2,424. Comparing the average AIC between models, we observe a slight decrease from 5,127 to 4,856.

While metric improvements are observed, the initial near-zero median residuals suggest that the basic model does not systematically mis-predict hotspots. Turnover's inclusion in the logistic regression does not drastically enhance model fit but adds complexity to computation and deployment. Model 4 produces a single log-odds value for each industry-expenditure code combination. These log-odds are easily converted to probabilities, a full matrix of which is provided in Supplementary Information SI5. Figure 3 displays an aggregated version of this matrix, plotting the most likely hotspot heading occurring by SIC Section.

Even at this aggregate level, we observe a complex web of hotspots. We observe frequent, industry-specific hotspots, where firm SIC section aligns closely with the spend category. For example, Section Q. Human Health and Social Work Activities exhibits a hotspot in Health-related spending. Supplementary Information SI5 shows that this is largely driven by SIC codes 86. Human Health Activities and 87. Residential Care Activities), with corresponding hotspots in categories such as 296 (Medical Facilities & Services) and 306 (Assisted Living and Long-Term Care). Similar patterns emerge across other industries, suggesting that a substantial share of firms' Scope 3 emissions arises from intra-industry transactions. This aligns with supply and use table patterns, where the largest values occur along the leading diagonal, reflecting industries' specialisation in their primary products (Australian Bureau of Statistics, 2021).

In addition, several expenditure codes consistently emerge as hotspots across multiple SIC sections. The General Business & Industry heading is particularly prominent, containing hotspots for 11 of the 14 modelled sections, while Food & Drink and Business Travel contain hotspots in five and four sections, respectively. Supplementary Information SI5 shows that General Business & Industry encompasses a diverse set of spend categories that frequently return high hotspot probabilities, especially among industrial SIC codes such as manufacturing and construction. Within this heading, Category 132 (Office Supplies) is the most common hotspot, appearing in 27 of the 57 SIC divisions modelled. Other recurring hotspots include Shopping (27 divisions) and Food & Drink (19 divisions).

4 Discussion

4.1 Key findings

The FTD-based approach of this paper offers new insights into the drivers of firms' Scope 3 emissions and facilitates analysis of the extent to which these patterns differ across industries. We identify that across all industries, a small number of spend categories typically account for an average of 75% of a firms calculated Scope 3 emissions. These categories are largely industry-specific, with a median of six per industry, while individual firms within those industries typically exhibit more concentrated spending patterns, with a median of two per firm.

We then construct models to predict these emissions, a practice often used by emission data providers (Serafeim & Velez Caicedo, 2022). We find that with just industry and turnover inputs; a hierarchical log–log regression specification achieves an RSQ of 0.87 for Category 1. Purchased Goods and Services. In this model, we observe a strong and significant relationship between turnover and Category 1 emissions with an elasticity of 0.87. The strong explanatory power of this simple model reflects both the standardized process of emissions calculation and the relatively coarse view it provides of emissions, whilst remaining consistent with previous studies showing that firm size and industry are primary determinants of Scope 3 emissions (Buchenau et al., 2025).

For Category 6. Business Travel and Category 9. Downstream Transportation and Distribution, model performance is weaker, with the simplest specifications returning RSQ values of 0.55 and 0.41, respectively. Whilst more complex iterations offering little improvement, these categories rely on stronger assumptions, which may contribute to the unexplained variance. For Business Travel, we assume all employee travel is paid directly via the business account, excluding cases where employees pay and later invoice the firm. For Downstream Transportation, many firms (43% of the sample) record zero spending in the relevant categories, further limiting model applicability. With these categories, emissions may reflect firm behaviours that extend beyond size and industry alone.

In the case of Category 3. Fuel- and energy-related, we find that the models developed in Phillpotts et al. (2025), can be redeployed as well-to-tank alternatives. The results are consistent with those reported in the paper, with the same model specifications yielding RSQ values of 0.87 and 0.72 respectively and a median average percentage error of 49%.

Our hotspot predictions complement the emission estimates by identifying spend categories that contribute disproportionately to calculated Scope 3 emissions. We observe both industry-specific hotspots, closely tied to sectoral activities, and shared categories that recur across multiple industries. These findings may indicate that many Scope 3 emissions originate within close supplier relationships, while also showing that certain categories of spending consistently drive emissions across sectors.

4.2 Practical implications

Mainstreaming the measurement of Scope 3 emissions remains highly challenging due to the complexity of measurement and the absence of harmonised reporting requirements. As a result, current Scope 3 data is often incomplete, and existing measurements are regularly inconsistent (Nguyen et al., 2023). This paper examines the potential of FTD as a transparent and standardised basis for calculating Scope 3 emissions, minimising subjective decision-making and reducing common challenges such as truncation error (Ward et al., 2018). For SMEs, which face financial, resource, and influence obstacles to Scope 3 measurement, the advantages of a low-cost, time-efficient solution are pronounced (Serafeim & Velez Caicedo, 2022).

Recent studies on household emissions (Trendl et al., 2023; Wells et al., 2025) highlight the FTD approaches, and our findings extend these insights to business emission accounting. For Scope 3 emissions the advantages of FTD are particularly salient given current regulatory gaps and limited awareness. Drawing on intrinsic business data that directly reflects operational behaviours, FTD offers a simplified pathway for SMEs to begin understanding and managing their supply chain impacts, which typically account for much of their environmental footprint.

To demonstrate the practical benefits of FTD, we use detailed, micro-level data to construct macro-level pathways that substitute for transaction records where these are unavailable. With only two user inputs, we generate robust estimates for Category 1. Purchased Goods & Services and Category 3. Fuel and Energy-related. These estimates can be refined by identifying recurring emission hotspots that arise from typical industry spending patterns.

To approximate total Scope 3 emissions, one accurately measured category can be scaled using benchmarking proportions to estimate the remaining categories (Buchenau et al., 2025). This approach constructs a complete picture of Scope 3 emissions, while avoiding reliance on the weaker predictions associated with Categories 6 and 9. An illustration of this approach is provided in Supplementary Information SI6.

When combined with Scope 1 and 2 estimates (Phillpotts et al., 2025), a simple model requiring only two user inputs produces a complete SME emissions profile. Guidance on accessing the model is provided in Supplementary Information SI6. Providing such estimates address the structural barriers faced by SMEs (Caldera et al., 2019; Menon & Ravi, 2021) and fosters broader engagement with emissions profile understanding among firms not yet subject to mandatory disclosure requirements (BBB, 2024). Understanding Scope 3 emissions enables businesses to identify their most material value-chain hotspots. Upstream reduction levers may include supplier engagement and procurement reform, with evidence indicating greater supply-chain engagement encourages more companies to disclose their emissions through voluntary frameworks (Science Based Targets initiative, 2018).

For policymakers, the findings highlight the value of leveraging FTD held by institutions to generate automated, standardised emissions insights. While financial institutions are increasingly aware of this potential (e.g., Bankers for Net Zero, 2022), policy guidance and incentives remain necessary to facilitate widespread, standardised deployment. Policymakers therefore have a role in creating the conditions that allow transaction-based approaches to complement existing reporting systems. Furthermore, our analysis also shows that upstream emissions are highly concentrated and overlapping across industries, revealing shared hotspots that can act as policy leverage points (Meadows, 1999), where modest but well-coordinated interventions could materially reduce overall emissions (Science Based Targets initiative, 2018).

4.3 Limitations

While these modelling solutions offer advantages in terms of accessibility and efficiency, it is important to acknowledge some key limitations. Our study includes 57 of the 88 SIC codes. To assess the impact of exclusions, Table S2 in Supplementary Information SI1 compares business population data (BEIS, 2022) with model coverage, showing that the model captures over 75% of the UK SME population. Among excluded SMEs, agriculture accounts for 5% and service sectors for 16%, both omitted due to inherent limitations in the FTD process.

Additionally, FTD lack the level of detail available in receipt-level data (Trendl et al., 2023). Emission conversion factors are therefore applied at the merchant category level only, making it difficult to maximise interpretation from hotspot estimates. For example, the general "Shopping" spend classification, would require examples of relevant merchants to assist transferability.

Related to this point, the validity of FTD-based Scope 3 emission estimates is difficult to assess, owing to accessibility and inconsistency limitations in alternative data. This challenge is compounded by the mapping of EEMRIO conversion factors to spend categories, which are derived from broad industry averages and applied to non–product-specific expenditures, exact matches are often unattainable. Instead, they serve only to differentiate between higher- and lower-embodied emission categories.

In practice, analysts of Scope 3 emissions should treat both reported and estimated values with caution and account for potential data error in their analyses (Nguyen et al., 2023). With an approach based on FTD, transparent and standardised methodologies come at the expense of absolute accuracy.

Here we demonstrate that some Scope 3 categories are either entirely unrepresented by FTD or require assumptions for compatibility (Table 1). This means that potentially significant emissions like Use of Sold Products are unmeasured and thus require their own measurement process, whilst those that require assumptions can be measured with an unknown degree of uncertainty. To address this, we propose the use of benchmarking approaches; however, these techniques rely on self-reported emissions data, which are often limited, incomplete, and inconsistent.

Finally, in addition to the limitations of a transaction approach, this methodology shares the limitations experienced by any IOA method. These include the time lags of EEMRIO databases, sectoral aggregations, price; temporal and spatial variation, and proportionality assumptions (Lenzen, 2000).

4.4 Future research directions

Among the limitations discussed, two key directions for future research emerge. The first refers to the limitations of EEMRIO emission factors. While these multipliers capture upstream emissions associated with each unit of expenditure, their publication is subject to significant time lags. This delay introduces inaccuracies, as economic structures and prices may change during the lag period. In the UK and across much of Europe, this issue is especially relevant due to recent experiences with prolonged periods of high inflation. A significant benefit to FTD is its real-time nature, and lagged multipliers depreciate this. To mitigate this challenge, techniques such as nowcasting and inflation adjustments have been proposed (OECD, 2017; Stadler et al., 2018). However, there remains a notable gap in the literature evaluating the effectiveness of these approaches in reducing inaccuracies associated with time lags.

A second avenue for future research is to identify the most suitable financial data for SME emission accounting. While this study focuses on FTD, accountancy-level data offers an alternative view of business operations. Its established role in financial reporting and subsequent completeness could reduce reliance on assumptions whilst improving coverage and accuracy. Similarly, the validity of FTD-based estimates should be tested against survey or physical data, to strengthen its use case in policy. Trendl et al. (2023) demonstrates the strong alignment of FTD with survey data across UK households; a comparable assessment for Scope 3 would be valuable, though currently constrained by data availability.

5 Conclusion

This paper demonstrates the potential of FTD as a scalable and transparent methodology for calculating key categories of Scope 3 emissions thereby addressing an important gap in current measurement practices. Using FTD-based estimates, we find emission sources are typically concentrated in a small set of hotspot spend categories, with both persistent industry-specific hotspots and common hotspots identified across multiple sectors. The models produced here robustly estimate Category 1 and Category 3 Scope 3 emissions, whilst Categories 6 and 9 are less well captured, reflecting their reliance on stronger assumptions.

FTD-based approaches substantially reduce reporting burdens and dependence on subjective measurement decisions, providing an accessible entry point for SMEs. Using the approach outlined here, SMEs can begin to explore decarbonisation strategies that extend beyond their direct operations, whilst the close relationships observed between firms Scope 3 emission sources highlight opportunities for coordinated, high-impact policy interventions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s44498-026-00003-5.

Acknowledgements We would first like to thank Lara Vomfell, Jasmine Wells, Alistair Gilfillan, and the rest of the behavioural science team at Lloyds Banking Group for sharing their invaluable expertise at key points of this research, as well as their continued interest and support. Secondly, we thank Hamza Suleman, whose contributions and technical skills were instrumental in developing the model described here into a working and accessible implementation. Finally, we extend our sincere gratitude to Ranil Boteju, Trystan Davies, Andrew McCulloch, Neil Oliver and Jaideep Mann for making this project possible and supporting its successful publication.

Author contributions The conceptualisation of the paper was carried out by AP and DL. AP conducted all analyses, with support from AT. AP drafted the manuscript, which was further refined through feedback and suggestions from AO, JN, AT and DL. JG provided technical guidance and support, and NJ oversaw and facilitated the overall project.

Funding Alec Phillpotts reports financial support was provided by Lloyds Banking Group. Jonathan Norman's time was funded by the UK Research and Innovation (UKRI) Council, as part of the Energy Demand Research Centre, grant reference number EP/Y010078/1.

Data availability The data that supports the findings of this study are available from Lloyds Banking Group, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data may be available from the authors upon reasonable request and with permission from LBG.

Declarations

Conflict of interest The authors declare no conflict of interest. The views and opinions expressed are those of the authors and do not necessarily reflect the views of Lloyds Banking Group (LBG) plc, its affiliates, or employees.

Ethical approval The LBG Privacy Risk and Impact Assessment process recognises the ethical basis for this study on aggregated, anonymous data as part of a strategy to help customers' transition to a more sustainable future. Upon opening an account, LBG customers consent to their data being used for research: https://www.lloydsbank.com/help-guidance/privacy/data-privacy-notice.html.

References

Acquaye, A., Genovese, A., Barrett, J., & Lenny Koh, S. C. (2014). Benchmarking carbon emissions performance in supply chains. *Supply Chain Management: An International Journal, 19*, 306–321. https://doi.org/10.1108/SCM-11-2013-0419

Ali, A., Ali, S., Khan, S. A., Khan, D. M., Abbas, K., Khalil, A., Manzoor, S., & Khalil, U. (2019). Sample size issues in multilevel logistic regression models. *PLoS ONE, 14*, Article e0225427. https://doi.org/10.1371/journal.pone.0225427

Australian Bureau of Statistics, 2021. Australian System of National Accounts: Concepts, Sources and Methods. p. Chapter 22. https://www.abs.gov.au/statistics/detailed-methodology-infor mation/concepts-sources-methods/australian-system-national-accounts-concepts-sources-and-methods/2020-21/chapter-22-input-output-tables (Accessed 25 Jun 2025).

Bankers for Net Zero, 2022. Scoping Exercise: The role of banks in reducing GHG emissions of UK SMEs. https://www.bankersfornetzero.co.uk/wp-content/uploads/2024/05/Smart-Data-Foundry-B4NZ-Scoping-Exercise-Full-report-September-2022-Final-version.pdf (Accessed 20 Aug 2025).

British Business Bank, 2024. Small Business Finance Markets Report 2024. https://www.british-business-bank.co.uk/about/research-and-publications/small-business-finance-markets-report-2024 (Accessed 25 Jun 2025).

Buchenau, N., Oetzel, J., & Hechelmann, R.-H. (2025). Category-specific benchmarking of Scope 3 emissions for corporate clusters. *Renewable and Sustainable Energy Reviews, 208*, Article 115019. https://doi.org/10.1016/j.rser.2024.115019

Busch, T., Johnson, M., & Pioch, T. (2022). Corporate carbon performance data: Quo vadis? *Journal of Industrial Ecology, 26*, 350–363. https://doi.org/10.1111/jiec.13008

Caldera, H. T. S., Desha, C., & Dawes, L. (2019). Evaluating the enablers and barriers for successful implementation of sustainable business practice in 'lean' SMEs. *Journal of Cleaner Production, 218*, 575–590. https://doi.org/10.1016/j.jclepro.2019.01.239

CDP, 2020. CDP Full GHG Emissions Dataset: Technical Annex IV: Scope 3 Overview and Modelling. https://assets.ctfassets.net/v7uy4j80khf8/4wN17Ev2pHtHqmqkDNvcA4/0ec997081f246af9a1d6d1bf69afbcd0/Technical_Annex_III_Scope_3_Overview_and_Modelling_2024.pdf Accessed 12 Jul 24.

Cheema-Fox, A., LaPerla, B. R., Serafeim, G., Turkington, D., & Wang, H (Stacie). (2021). Decarbonizing Everything. *Financial Analysts Journal, 77*, 93–108. https://doi.org/10.1080/0015198X.2021.1909943

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd ed.). Routledge.

Companies Act, 2006. Legislation.gov. https://www.legislation.gov.uk/ukpga/2006/46/contents (Accessed 12 Dec 2024).

Department for Business, Energy & Industrial Strategy, 2022. Business population estimates for the UK and regions 2022: statistical release . GOV.UK. https://www.gov.uk/government/statistics/business-population-estimates-2022/business-population-estimates-for-the-uk-and-regions-2022-statistical-release-html (accessed 7.12.23).

Department for Energy Security and Net Zero, 2023. UK greenhouse gas emissions reporting: Scope 3 emissions . GOV.UK. https://www.gov.uk/government/calls-for-evidence/uk-greenhouse-gas-emissions-reporting-scope-3-emissions Accessed 1 Jul 2025.

EU, 2022. Directive (EU) 2022/2464 of the European Parliament and of the Council, OJ L. https://eur-lex.europa.eu/eli/dir/2022/2464/oj/eng Accessed 11 Aug 2025.

Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. *Higher Education from Cambridge University Press*. https://doi.org/10.1017/CBO9780511790942

Goldhammer, B., Busse, C., & Busch, T. (2017). Estimating corporate carbon footprints with externally available data. *Journal of Industrial Ecology, 21*, 1165–1179. https://doi.org/10.1111/jiec.12522

Hausman, J. (2001). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives, 15*, 57–67. https://doi.org/10.1257/jep.15.4.57

Hertwich, E. G., & Wood, R. (2018). The growing importance of scope 3 greenhouse gas emissions from industry. *Environmental Research Letters, 13*, Article 104013. https://doi.org/10.1088/1748-9326/aae19a

Hettler, M., & Graf-Vlachy, L. (2024). Corporate scope 3 carbon emission reporting as an enabler of supply chain decarbonization: A systematic review and comprehensive research agenda. *Business Strategy and the Environment, 33*, 263–282. https://doi.org/10.1002/bse.3486

Huang, Y., Lenzen, M., Weber, C., Murray, J., & Matthews, H. (2009). The role of input-output analysis for the screening of corporate carbon footprints. *Economic Systems Research, 21*, 217–242. https://doi.org/10.1080/09535310903541348

Isil, O., & Sebastianelli, R. (2020). Arcs of carbon awareness in the value chain and their antecedents. *Business Strategy and the Environment, 29*, 503–518. https://doi.org/10.1002/bse.2383

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning with Applications in R* (2nd ed.). Springer.

Kitzes, J. (2013). An introduction to environmentally-extended input-output analysis. *Resources, 2*, 489–503. https://doi.org/10.3390/resources2040489

Lenzen, M. (2000). Errors in conventional and input-output-based life-cycle inventories. *Journal of Industrial Ecology, 4*, 127–148. https://doi.org/10.1162/10881980052541981

Lloyds Banking Group (LBG), 2023. 2023 Lloyds Banking Group Annual Review. https://www.lloydsbankinggroup.com/assets/pdfs/investors/annual-report/2023/2023-lbg-annual-review.pdf Accessed 25 Jun 2025.

Meadows, D.H., 1999. Leverage Points: Places to Intervene in a System. The Sustainability Institute. https://donellameadows.org/archives/leverage-points-places-to-intervene-in-a-system/ Accessed 23 Aug 2025.

Menon, R. R., & Ravi, V. (2021). Analysis of barriers of sustainable supply chain management in electronics industry: An interpretive structural modelling approach. *Cleaner and Responsible Consumption, 3*, Article 100026. https://doi.org/10.1016/j.clrc.2021.100026

Minx, J., Wiedmann, T., Barrett, J., Suh, S., 2008. Methods review to support the PAS for the calculation of the embodied greenhouse gas emissions of goods and services. https://www.academia.edu/13468787/Methods_review_to_support_the_PAS_process_for_the_calculation_of_the_greenhouse_gas_emissions_embodied_in_goods_and_services Accessed 23 May 2025.

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*, Article e0000208. https://doi.org/10.1371/journal.pclm.0000208

OECD, 2017. Nowcasting techniques to improve timeliness. Presented at the Committee for the Coordination of Statistical Activities, Muscat.

Office for National Statistics (ONS), 2022. Supply and use tables. https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables (Accessed 8 Apr 2023.

Owen, A., Kilian, L., 2024. 2024 Data Release of Consumption- based Accounts for the UK: Summary of Methods. https://assets.publishing.service.gov.uk/media/68220ff3db6463b14cd81916/Summary_of_Methods_2025_final.pdf Accessed 01 Sept 2025

Patchell, J. (2018). Can the implications of the GHG protocol's scope 3 standard be realized? *Journal of Cleaner Production, 185*, 941–958. https://doi.org/10.1016/j.jclepro.2018.03.003

Phillpotts, A., Owen, A., Norman, J., Trendl, A., Gathergood, J., Jobst, N., & Leake, D. (2025). Bridging the SME reporting gap: A new model for predicting scope 1 and 2 emissions. *Journal of Industrial Ecology*. https://doi.org/10.1111/jiec.70106

Robinson, O., Kemp, S., & Williams, I. (2015). Carbon management at universities: A reality check. *Journal of Cleaner Production, 106*, 109–118. https://doi.org/10.1016/j.jclepro.2014.06.095

Schmidt, M., Nill, M., & Scholz, J. (2022). Determining the scope 3 emissions of companies. *Chemical Engineering & Technology, 45*, 1218–1230. https://doi.org/10.1002/ceat.202200181

Science Based Targets initiative, 2018. Value Change in the Value Chain: Best Practices in Scope 3 Greenhouse Gas Management. https://files.sciencebasedtargets.org/production/files/SBT_Value_Chain_Report-1.pdf Accessed 07 Oct 2025.

Serafeim, G., Velez Caicedo, G., 2022. Machine Learning Models for Prediction of Scope 3 Carbon Emissions. https://doi.org/10.2139/ssrn.4149874

Shakdwipee, M., Lee, L.-E., 2016. Filling the Blanks: Comparing Carbon Estimates Against Disclosures (MSCI ESG Research Issue Brief). MSCI.

Stadler, K., Wood, R., Bulavskaya, T., Sodersten, C.-J., Simas, M., Schmidt, S., Usubiaga, A., Acosta-Fernandez, J., Kuenen, J., Bruckner, M., Giljum, S., Lutter, S., Merciai, S., Theurl, M. C., Plutzar, C., Kastner, T., Eisenmenger, N., Erb, K.-H., Koning, A., & Tukker, A. (2018). EXIOBASE 3 developing a time series of detailed environmentally extended multi-regional input-output tables. *Journal of Industrial Ecology, 22*, 502–515. https://doi.org/10.1111/jiec.12715

Trendl, A., Owen, A., Vomfell, L., Kilian, L., Gathergood, J., Stewart, N., & Leake, D. (2023). Estimating carbon footprints from large scale financial transaction data. *Journal of Industrial Ecology, 27*, 56–70. https://doi.org/10.1111/jiec.13351

Tukker, A., de Koning, A., Owen, A., Lutter, S., Bruckner, M., Giljum, S., Stadler, K., Wood, R., & Hoekstra, R. (2018). Towards robust, authoritative assessments of environmental impacts embodied in trade. *Journal of Industrial Ecology, 22*, 585–598. https://doi.org/10.1111/jiec.12716

Ward, H., Wenz, L., Steckel, J. C., & Minx, J. C. (2018). Truncation error estimates in process life cycle assessment using input-output analysis. *Journal of Industrial Ecology, 22*, 1080–1091. https://doi.org/10.1111/jiec.12655

Wells, J., Trendl, A., Owen, A., Barrett, J., Jobst, N., & Leake, D. (2025). Targeting carbon reduction in UK households: A new segmentation model using financial transaction data. *Journal of Industrial Ecology*. https://doi.org/10.1111/jiec.70020

Wiedmann, T., Wilting, H. C., Lenzen, M., Lutter, S., & Palm, V. (2011). Quo vadis MRIO? Methodological, data and institutional requirements for multi-region input–output analysis. *Ecological Economics, 70*, 1937–1945. https://doi.org/10.1016/j.ecolecon.2011.06.014

World Business Council for Sustainable Development and World Resources Institute (WRI and WBCSD), (2020). Corporate Value Chain (Scope 3) Accounting and Reporting Standard. World Resources Institute and World Business Council for Sustainable Development. https://ghgprotocol.org/sites/default/files/standards/Corporate-Value-Chain-Accounting-Reporing-Standard_041613_2.pdf. Accessed 6 June 2025