Analysing community-level spending behaviour contributing to high carbon emissions using stochastic block models

Keywords: financial networks, spending behaviour, carbon emissions, green economy, community detection

Extended Abstract

Reducing carbon emissions is a central objective for governments and institutions worldwide. Many countries have committed to reaching net zero by mid-century, and sustainability targets increasingly emphasise the role of household consumption in this process. Financial institutions responded by developing tools linking spending to emissions, and recent open approaches show that financial transaction data, combined with official expenditure and emissions statistics, provides a reliable way of estimating consumption-related carbon footprints [1].

A persistent challenge is translating individual-level data into group-level insights that support scalable interventions. Segmenting consumers into communities with shared spending and emissions patterns allows for the design of practical and targeted policies. However, there is limited research analysing debit card transaction datasets and the emissions associated with spending. Most existing studies that use transaction data segment consumers based on emissions per category and socio-economic or demographic attributes [2], which risks reinforcing existing inequalities and does not capture structural patterns in the emissions resulting from different spending profiles.

To address this, we obtain a debit card transaction dataset from ekko, a sustainable banking FinTech, containing 52,496 transactions from 1,362 UK customers between 2021 and 2023. We construct a bipartite network of customers and merchant categories (see Figure 1a), initially with unweighted links, indicating whether a customer makes a transaction in a given category. We then extend this network by weighting each link according to spending and the associated carbon emissions.

We apply a bipartite Stochastic Block Model (SBM) community detection algorithm [3] (see Figure 1b) to both networks. The unweighted SBM clusters consumers with similar spending patterns across categories based only on whether they have transactions in each category, without considering the magnitude of spending or emissions. In contrast, the weighted SBM incorporates both spending and the associated carbon emissions for each customer–category link, producing communities where the average emissions per category are comparable across clusters. By comparing unweighted and weighted SBM results, we can therefore identify communities based on either shared transaction behaviour or cluster-level emissions equivalence, revealing different structural aspects of the network.

Community-level analysis demonstrates that the network-based approach uncovers heterogeneous patterns of behaviour and emissions. Some communities are dominated by high-emission categories such as private transport or gas service stations, while others concentrate in lower-emission domains such as public transport or retail. A small number of communities contribute disproportionately to total emissions due to concentrated activity in specific categories. These insights illustrate the value of network-based modelling for identifying influential groups and informing interventions targeting collective behavioural patterns rather than individual transactions.

Financial transaction networks can inform interventions to reduce emissions at the community level. By identifying communities with similar spending or emissions profiles, policymakers and financial institutions can target initiatives such as subsidies for low-carbon alternatives or tailored communication for the most impactful groups. Group-level interventions are particularly important because influencing behaviour individually is neither scalable nor practical.

This research underwent full ethics review and approval, and all analyses are conducted in accordance with established ethics guidelines. All data are anonymised before analysis, and no personally identifying information is used. From the start, we discussed potential ethical risks with both ekko, the data owner, and our academic institution, and took steps to reduce these risks, such as preventing re-identification of individuals, guiding policy recommendations to remain fair and equitable, and acknowledging data limitations that could affect the interpretation or application of the results. These discussions informed decisions on data handling, aggregation, and analysis, helping to protect customer privacy while allowing meaningful research. We recognise that our findings may influence policy, and we provide guidance on designing such policies to be inclusive, fair, and socially responsible. We also considered the broader societal implications of our work and emphasised responsible communication to support its ethical use.

References

[1] Anna Trendl et al. "Estimating carbon footprints from large scale financial transaction data". In: *Journal of Industrial Ecology* 27.1 (2023), pp. 56–70.

[2] Jasmine Wells et al. "Targeting carbon reduction in UK households: A new segmentation model using financial transaction data". In: *Journal of Industrial Ecology* (2025).

[3] Martin Gerlach, Tiago P Peixoto, and Eduardo G Altmann. "A network approach to topic models". In: *Science Advances* 4.7 (2018), eaaq1360.

(a) An example of a transaction bipartite network, showing the connection between customers (top) and merchant categories (bottom).
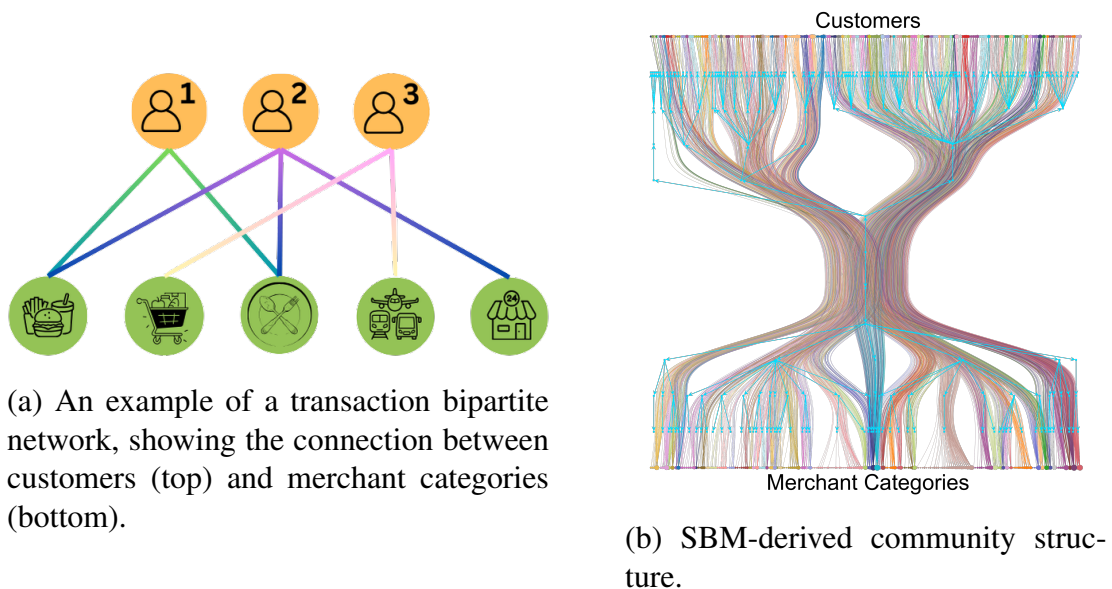

(b) SBM-derived community structure.

Figure 1: Illustration of customer spending communities. Panel (a) shows a simple representation of customer transactions across merchant categories, and panel (b) shows Stochastic Block Model clustering revealing communities with similar carbon-linked spending.