Imperial College London

Department of Medicine - School of Public Health

Construction of an Environmental Behavioural Score from Visa Transactional Data

Author:
02112494

Word Count:
9998

Supervisors:
Prof Marc Chadeau-Hyam
Dr Ana Pengelly
Dr Gianluca Campanella

A thesis submitted for the degree of

MSc Health Data Analytics and Machine Learning

December 10, 2024

Acknowledgements

I want to thank my supervisors and the Visa Consulting & Analytics Data Science Team for ongoing advice and support.

Abstract

Introduction

The climate crisis is becoming increasingly problematic; global carbon emissions must be reduced to prevent severe harm to our ecosystem. This exploratory study investigates whether transactions can be used to score individuals and local authority districts (LADs) in London for their environmental behaviours during the first six months of 2022. Further analysis is carried out to break down these behaviours and better understand their role in individuals' carbon emissions. A secondary analysis was conducted to see how scores and behaviours correlate with asthma prevalence across London.

Methods

The analysis was conducted in three steps. Step one was to identify environmental behaviours and extract their corresponding transactions. Step two was calculating a coefficient for the impact of each behaviour based on literature. Step three was calculating the score with the established coefficients after determining and adjusting for baseline behaviours. This score was then compared with the domestic carbon emissions by LAD as well as asthma prevalence by LAD. A sensitivity test was conducted to assess the correlation between score and all emissions produced within the LADs.

Results

The results show a significant positive correlation between the score and domestic carbon emissions by LAD. A few specific behaviours were identified to lead to the most emissions at the LAD level: gas vehicle ownership and operation, flights, and clothing purchases. At the individual level, it can be observed that the operation of an electric or hybrid vehicle can help reduce emissions due to driving; however, these individuals are associated with higher spending on flights and higher emissions overall.

A significant positive correlation exists between the score and asthma. Positive correlations also existed between asthma prevalence and gas vehicle operation. A negative correlation exists between asthma prevalence and flying, as well as asthma prevalence and electric vehicle usage. This relationship is likely confounded by wealth.

A correlation existed between many of the behaviours, implying that the more carbon emitted through one behaviour, the more an individual is expected to emit via other behaviours; as a direct consequence, the score correlates with total spending at the PAN level.

London LAD Key

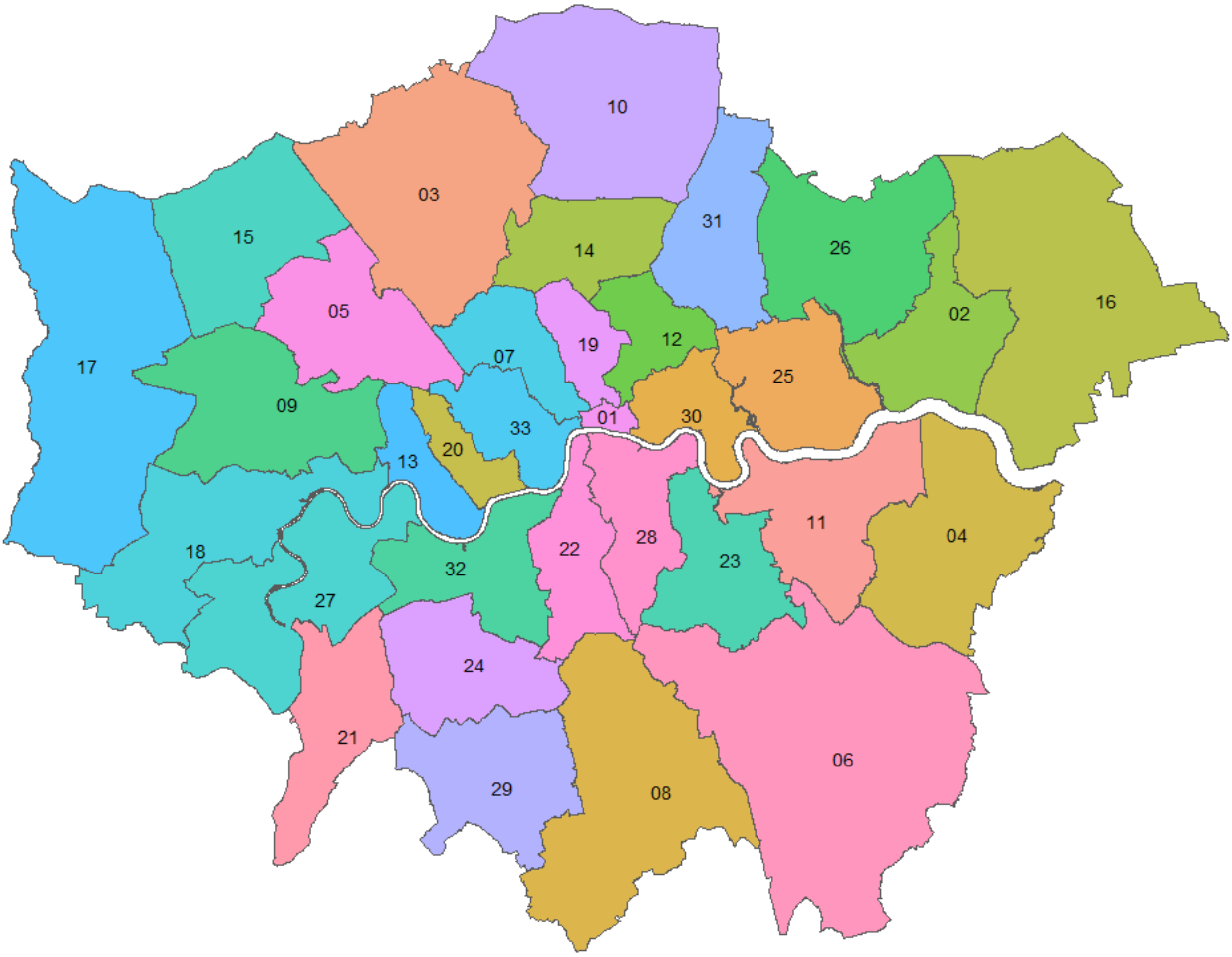


Figure 1: LADs and their final two digits, E090000%%

E09000001 = City of London
E09000002 = Barking and Dagenham
E09000003 = Barnet
E09000004 = Bexley
E09000005 = Brent
E09000006 = Bromley
E09000007 = Camden
E09000008 = Croydon
E09000009 = Ealing
E09000010 = Enfield
E09000011 = Greenwich
E09000012 = Hackney
E09000013 = Hammersmith and Fulham
E09000014 = Haringey
E09000015 = Harrow
E09000016 = Havering
E09000017 = Hillingdon
E09000018 = Hounslow
E09000019 = Islington
E09000020 = Kensington and Chelsea
E09000021 = Kingston upon Thames
E09000022 = Lambeth
E09000023 = Lewisham
E09000024 = Merton
E09000025 = Newham
E09000026 = Redbridge
E09000027 = Richmond upon Thames
E09000028 = Southwark
E09000029 = Sutton
E09000030 = Tower Hamlets
E09000031 = Waltham Forest
E09000032 = Wandsworth
E09000033 = Westminster

Contents

List of Abbreviations

EV: Electric Vehicle including electric cars and hybrid cars (those with electric capabilities)

LAD: Local Authority District

MCC: Merchant Category Code

MCG: Merchant Category Group

PAN: Primary Account Number: 16 unique digits representing each card

PI: Personal Information

POS: Point of Sale

VCA: Visa Consulting & Analytics

List of Figures

List of Tables

Introduction

1.1 Background

So far, global warming has reached approximately 1°C above pre-industrial levels, and this number continues to rise at around 0.2°C per decade. Under The Paris Agreement of 2015, 194 countries agreed to reduce their emissions, to cap climate change at 2°C [3]. At its current level, climate change has caused more extreme weather conditions worldwide. Global warming has impacted poorer nations the most, risking the security of food and drinking water; soon, parts of the planet will be uninhabitable [4]. Global warming impacts more prosperous nations, too. The UK has seen an increase in the rates of floods and droughts [5]. In July 2022, the UK recorded its hottest temperature since records began of 40.3°C [6], leading to fires, failing infrastructure, and 13 recorded deaths [7]. In 2018, the Intergovernmental Panel on Climate Change (IPCC) warned of the disastrous consequences if global warming surpassed 1.5°C. To avoid surpassing this level, the IPCC have said CO2 emissions must peak in or before 2025 [8].

Figure 1.2: Contributions to the average individual's carbon footprint, per year. [1]

Figure 1.2 shows the breakdown of the average person's carbon footprint in the UK as of 2011. There are five main categories. The most significant component is an individual's purchases; this includes clothes, household goods and entertainment. Fast fashion continues to become a more significant contributor to emissions, with the production of clothes doubling from 2000 to 2015 [9]; the United Nations (UN) Environmental Project have now estimated that fashion accounts for 10% of call carbon emissions [10]. Contributing 2.9tCO2 to an individual's yearly carbon footprint is food and drink. Red meats are the main culprit within this category, followed by dairy. Only 5% of food and drink emissions are as a result of transporting goods [11]. The third largest category is Home Energy which contributes 2.5t CO2 per person per year; this category comprises the carbon produced from the fuel used to power and heat homes. More efficient housing can use as little as 0.5tCO2 per person-year, while less efficient housing can be 4tCO2. With 1.9tCO2, the next category is travel. Travel emissions are highly variable per person; car usage and aviation are the most prominent two contributors to travel emissions. The last category is infrastructure, with total emissions of 1.7tCO2 per person per year; these result from government spending on public infrastructure such as schools and hospitals for the population. [1]

1.2 Prior Work and Gaps in the Literature

Several companies have created systems to estimate an individual's carbon footprint from their transactional data. One of these companies is MasterCard [12]. MasterCard uses a service named Doconomy, which breaks transactions down into subgroups, and assigns an impact to each subgroup; combining this information gives a weighted sum of the transactions, corresponding to the total carbon emissions attributed to an individual's purchases. Visa uses a similar system internally too. While this method accurately estimates carbon emissions, it does not provide much information on an individual's choices and these choices' effects on the environment. This method also heavily correlates with spending; with few spending habits reducing an individual's carbon footprint, higher spending will almost always increase an individual's carbon footprint.

No academic work has been conducted using transactional data to map behaviours against carbon emissions, nor has any work been conducted to score individuals from their environmental behaviours. Lack of research, in part, can be attributed to climate research being against financial institutions' interests since the general finding is that consumers should be spending less.

1.3 Research Motivation and Aims

The aim of this research project is twofold.

1. Construct an environmental behavioural score at the primary account number (PAN) and local authority district (LAD) level.
2. Further investigate behaviours and their environmental impacts.

To measure environmental behaviours, only carbon emissions are considered. A three-step process was followed to undertake these aims. Firstly, behaviours were established and mapped from the transactional data, which gave each PAN a set of behaviours. Secondly, the magnitude of these behaviours were found in the literature. Lastly, these behaviours and magnitudes were combined to form a score at the PAN level. An LAD level score was also generated by taking the mean score within each LAD. From here, the score was analysed, along with its ecological correlations to domestic CO_2 emissions and asthma prevalences, by LAD. Further analysis was conducted better to understand consumer choices and the carbon impact of each behaviour.

1.4 Transactional Data

For this research, the author was granted access to VisaNet data. VisaNet captures all transactions that appear with a Visa card, generating a vast quantity of naturally occurring data. It allows for unprecedented insights into consumer behaviour and economic change. Using VisaNet allows for a new approach to understanding individuals' behaviours toward the environment.

The anonymity of cardholders is of the utmost importance to Visa. PANs are encrypted to a 'random' string of 16 characters which the author cannot unencrypt. The only information associated with each PAN is whether it is a business card or an individual card and whether it is a debit or credit card; no address, behaviour, age, ethnicity or other personal information (PI) is stored.

1.5 Challenges

Transactional data comes with several limitations that are important to highlight prior to understanding the process followed. Transactional data comes at the PAN level; in 2020, it was documented that the average individual in the UK has 3.12 payment cards. [13]. These cards cannot be linked together since the PANs are independent, and no PI is known.

While an individual's purchases may be observed and corresponding behaviours inferred (i.e. if you purchase Mcdonald's, you can infer fast food consumption), it is virtually impossible to infer behaviour from a lack of observations. There are several justifications for an action not being observed: The action may not have happened, the action may not have a purchase associated with it (e.g. walking to work), or the action could be performed on another of the individual's cards, with cash, or it could have been purchased by someone else for the individual.

Methods

2.1 Data and Pre-processing

2.1.1 VisaNet

VisaNet works via a multi-step process. Before transactions can occur, Issuers (generally banks) give cards with Visa chips to individuals, and merchants accept Visa to process their payments. Then, when a cardholder makes a purchase at a merchant, the merchant gives the transaction information to the Acquirer (the merchant bank), who then passes the information to Visa. Visa pass the transaction information to the Issuer along with the addition of several metrics calculated by Visa, such as the likelihood that the transaction is fraudulent. The Issuer decides whether to accept or decline the transaction (if the Issuer is unavailable to verify a transaction, Visa stand-in to perform this step). Visa relay the verification outcome to the Acquirer, who relays the outcome to the merchant, who lets the cardholder know if their transaction was accepted at the point of sale [14]. Visa records all transaction requests and clearings between acquires and issuers. Only approved transaction requests are within the scope of this study. In total, Visa has around 3.8 billion active cards worldwide and is accepted at over 100 million merchants [15]. This uptake translates to visa processing 150 million transactions a day, with the capability of processing 24,000 transactions per second [16]. Each of these transactions is tabulated with several hundred data fields.

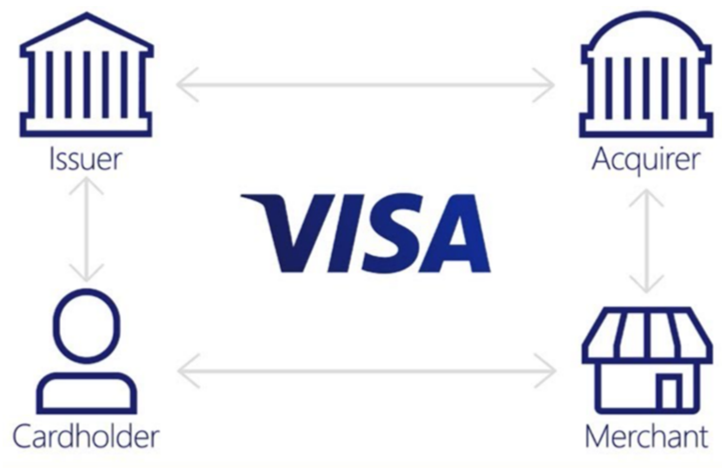


Figure 2.3: Visa Four Party System [2]

As a direct consequence of this system, the number of cards within the Visa Network is unknown since cards are issued by the Issuer, not by Visa. Visa is only aware of the cards that have used their payment network (generally by making a transaction).

Some filtering criteria were applied to obtain the dataset for this project. The study period is all those transactions between 1 January 2022 and 30 June 2022 (inclusive). Only transactions from debit cards registered in the United Kingdom were considered. Credit cards were rejected due to spending habits on credit cards being less informative due to the different ways in which people use credit cards. Some individuals use credit cards to build up a credit rating, so they will make small purchases and pay off the balance each month. In contrast, others depend on credit to make daily payments, with no ability to pay off the balance at the end of the month. Simply, credit card behaviours are more complex [17]. The second filter was for personal cards. Business cards follow a different spending pattern to personal cards, and any behaviours linked with a business card are likely the result of company policy and HR as opposed to an individual's habits.

There are 118,627,773 corresponding cards in the UK. Only PANs from London were considered since different locations lead to different opportunity-costs of actions, which the literature suggested would be problematic [18]. For example, going to a shop in London is always quick and easily accessible via foot or public transport. However, in the countryside, public transport links might be lacking, so the only quick and easy way might be to drive. This disparity in opportunity-cost means that an individual who drives in London should have a higher score (meaning worse environmental behaviours) than an identical individual from the countryside, keeping everything else constant.

In order to only consider Londoners, each PAN needed to be assigned to an LAD. To assign PANs to an LAD, an algorithm was developed. The previous algorithm used within the Visa Consulting & Analytics (VCA) team was to assign each PAN to the LAD, where it transacts the most within the given period. The author improved the LAD classification system using Bayesian methods, assuring that the proportion of cards assigned to an LAD does not exceed the proportion of cards expected in that LAD, given its population. This method assumes that the number of cards carried is independent of the region of London where an individual resides.

The algorithm developed is as follows:

1. Count the number of transactions within each LAD performed by each PAN. For each PAN, rank the LADs in order of the most transactions performed to the least.
2. Remove any PAN that transacts less than five times in its highest ranking LAD.
3. Calculate the proportion of transactions performed in each LAD compared to the number of transactions in the top 5 LADs. These values will be used as weightings to decide which LAD to assign to each PAN.
4. Using population statistics from 2020 [19], calculate the proportion of cards expected within each LAD.
5. Assign all PANs to their top-ranking LAD.
6. If the LAD exceeded the proportion of cards expected, the PANs with the lowest weightings were removed until the proportion expected was not exceeded.
7. Repeat the previous two steps with unassigned PANs that transact at least five times in their second top-ranking LAD.
8. Any PANs unassigned to an LAD are rejected due to low confidence in their classification.

The method outlined above means that PANs are assigned to the region they transact the most, second most or none. If a PAN is not assigned to its highest ranking LAD, they are deemed less likely to be in that group than other PANs, as reflected by a lower weighting. This classification method appears to be more accurate than the original method, as reflected by the distribution of cards in Figure 2.4 and the opinion of those more senior within the VCA team.

Figure 2.4 demonstrates the distribution of individuals within London and the distribution of card assignments according to the original algorithm and the author's algorithm. For the full breakdown, see table 6.3 in the appendix. While the original algorithm performs moderately well, certain LADs are highly misrepresented. These included too many PANs assigned to Westminster, likely due to the number of shopping districts like Oxford Street, which is the busiest shopping area in Europe [20]. City of London also had too many PANs assigned; with only 8,000 residents, some of the 500,000 daily commuters [21] were likely mislabelled. In the aforementioned LADs, most PANS are probably mislabelled and thus will not represent the LAD well. The new method mimics the population reasonably well except for Barking and Dagenham, which still sees an improvement upon the original method, implying that the region is likely more accurately represented. Figure 2.4 was generated after all pre-processing steps (more mentioned below) were performed; hence some of the new classification proportions slightly exceed actual population proportions.

The original algorithm assigned 7,622,724 cards to a London-based LAD, while the new system

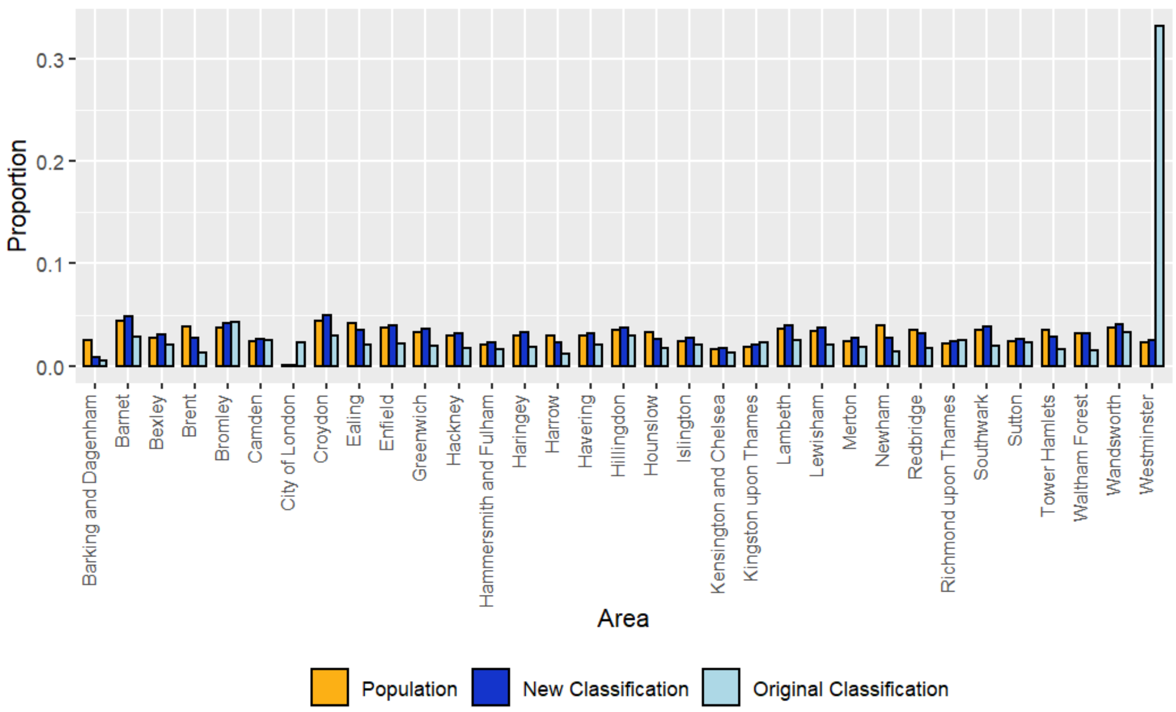


Figure 2.4: Proportion of PANS assigned to each LAD vs population proportions

assigned 4,649,861 PANS to a London LAD. Most of the 2,972,863 card discrepancy results from LAD clashes leading to rejected cards.

Due to individuals often using multiple cards and many PANs having low activity levels, a threshold number of transactions was set to deem a card 'active'. This threshold was 20 transactions a month within the considered period. Filtering for this threshold left 162,786 PANs remaining. Within the study period, these PANs performed 30,356,535 transactions cumulatively, which correspond to a total spend of £946,814,762.31.

The nature of transactional data is that it has a heavy positive skew. A minority of transactions are very large, and a minority of PANs have abnormal habits, such as an exceptionally high frequency of transactions in specific areas; these are potentially personal accounts being misused by businesses or are abnormal consumers. The behaviours identified were adjusted to account for this. If the behaviour was based on the number of transactions by each PAN, the 99th percentile number of transactions on that behaviour was identified. Any value above this was rounded down to equal the 99th percentile. When the behaviour was based on the total amount spent by each PAN, the 99th percentile amount spent on that behaviour was identified. Any value above this was rounded down to equal the 99th percentile. This method allowed extreme PANs to contribute to correlations without single PANs having too much influence.

2.1.2 Outcome Data

Several publicly accessible data sets were collected, lightly processed and used. The most recently available data was used. No significant changes are expected between the datasets being collected and the research period of this project.

LAD boundaries (2018 - present) for London were obtained from the London Datastore by the London Assembly in a Shapefile format [22]. The 2005 - 2019 domestic CO2 emissions by LAD [23] and the asthma prevalence by LAD (2019/20) [24] were also collected for the outcome variables. For the sensitivity test, total CO2 emissions in 2019 from each LAD were used; this includes CO2 from manufacturing, tourism, transport, and other sources [25].

2.2 Behaviours Identified

Upon reviewing the literature, behaviours and their respective impacts were identified. The net effect of each behaviour was calculated uniquely. Several coefficients were altered to account for baseline behaviours when constructing the environmental behavioural score. The impact of baseline behaviours is set to zero:

1. Local Travel: Since London has popular and extensive transport links, inner-city public transport (mainly TfL) was determined to be the baseline. The coefficients of gas for cars, electricity for cars, taxi usage and e-bike usage have been affected.

2. National Travel: The UK's train network is a popular and fast mode of transport between towns and cities. The coefficient for coach travel was altered accordingly.

3. Food: Individuals need food, so the coefficient is zero. This means the coefficient for food 'rescue' is negative.

Each transaction within VisaNet is automatically assigned to a merchant category code (MCC). For example, a train ticket from Trainline will be placed in the MCC 'Passenger Railways', while a Black Cab ride will be placed in 'TAXICABS'. Then multiple MCCs are assigned to a single merchant category group (MCG), so all transactions within 'Passenger Railways', 'TAXICABS', and dozens of others will come under the MCG of 'Public Transport'. There are a total of 30 MCGs and hundreds of MCCs. The categorisation for transactions relating to transport is relatively reliable since their quality was independently verified in 2019. The categorisation of other transactions is less reliable.

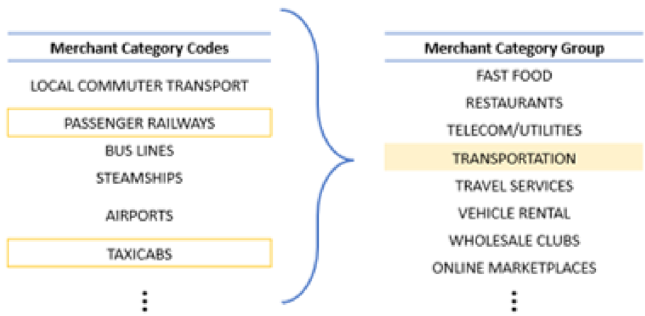


Figure 2.5: Categorisation of MCCs and MCGs

2.2.1 Air Miles

This category represents the emissions from air travel by the individual within the study period.

The emissions due to air miles were calculated using internal research at Visa, which concluded that there is an estimated 1000g/$ of CO2. Visa calculated this coefficient by using flight information provided to them, which gave access to departure and arrival airports for a subset of transactions; hence a coefficient could be calculated by comparing the price to the distance flown.

An MCG existed solely for transactions on air travel. The 100 merchants with the largest overall spend were screened; three mislabelled merchants were removed, and all other transaction within this MCG was used.

2.2.2 Gasoline

This category represents the emissions from the individual driving a gasoline-based vehicle.

The emissions were calculated using the gas sales price divided by the price per litre. The average amount of CO2 produced per litre of petrol is 2362g/L and for diesel is 2629g/L [26]. The price per litre was calculated by averaging the weekly average in the breakdown of fuel prices from gov.uk [27] giving petrol as 164.2p and diesel as 173.23p. The ratio of gas to diesel usage in the UK is 13:25 [28]. The average petrol vehicle in the UK performs 22.36km/L, and the average diesel vehicle performs 23.85KM/L [29].

These statistics give the total coefficient for each individual as

$$\frac{\pounds P}{38} \times \left(\frac{13 \times (2362 - \mathbb{1}(44 \times 22.36))}{1.642} + \frac{25 \times (2629 - \mathbb{1}(44 \times 23.85))}{1.7323} \right) := \begin{cases} 1 & \text{if adjusted for baseline,} \\ 0 & \text{if unadjusted.} \end{cases} \tag{2.1}$$

Where £P is the amount each individual purchased in pounds. This equation equates to 1491g/£ and 887g/£ after adjusting for the baseline of TfL.

An MCG exists for those who paid at the pump; all of these transactions were used. Another MCG exists for all purchases at the counter within gas stations. The distributions of pay-at-pump and pay-in-store transactions were analysed.

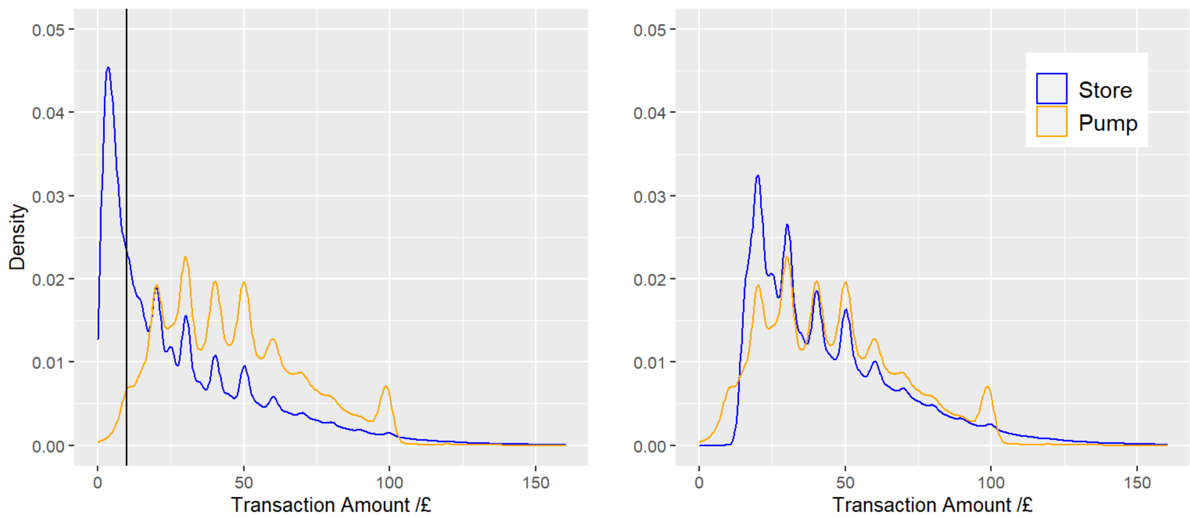


Figure 2.6: Fuel pump and store transaction distributions before and after threshold

Two density plots are shown in Figure 2.6, showing the distribution of transaction spending amounts. Paying at the pump shows very few transactions under £15. Several differing size peaks at £20, £30, £40, £50, and £60 can be observed, likely as a result of individuals topping their car up by round quantities. These peaks are followed by a smooth decreasing function with a small peak and an abrupt end at around £100. The abrupt end is likely due to payment caps at the pump [30]. Payments in-store show a decreasing function, with a peak around zero; this is likely due to individuals using petrol stations for convenience purchases such as lunch. A sharp decrease follows this up to around £15, then a gradual decrease till £60. Similar local maxima to the pay-at-pump transactions can be observed at the multiples of £10 between £20 and £60. Between £60 and £150, a smooth decreasing function can be seen with a shallow gradient. When removing in-store transactions below £15, the pay-at-pump and pay-in-store transactions trends are very similar. Thus, a £15 lower limit was set on in-store transactions for them to be deemed as fuel purchases to help remove non-fuel-based transactions.

2.2.3 Electricity (For Cars)

This category represents the emissions from individuals due to driving an electric-based vehicle.

The emissions were calculated by dividing the PAN's spend by the average cost per Kilowatt Hour to estimate how much electricity was purchased. Then the average emissions per kWh were used to determine the impact. Throughout the day, the carbon emissions from a unit of energy purchased vary according to current demand. Carbon emissions vary throughout the day because a higher proportion of energy can come from 'clean' forms such as nuclear or renewable when demand is low. However, the average carbon emissions are 221g/KWh [31] and the average price in May 2022 was 50.97p [32]. The average kWh per km is 0.2 [33].

$$\pounds P \times \frac{(221 - \mathbb{1} \cdot \frac{44}{0.2})}{0.5097} := \begin{cases} 1 & \text{if adjusted for baseline,} \\ 0 & \text{if unadjusted.} \end{cases} \tag{2.2}$$

Where £P is the amount each individual purchased in pounds. This equation equates to 434g/£ and 2g/£ after adjusting for the baseline

An MCG exists purely for electricity top-ups; it is tidy and can be used as it is.

2.2.4 Car Ownership

This category represents the carbon emissions attributable to an individual's vehicle manufacturing. Car manufacturing emits high quantities of carbon, especially in electric and hybrid (EV) vehicles, due to their batteries. As a result, the ownership of vehicles must be considered a behaviour, on top of fuel usage.

In order to determine car ownership, an extended period of transactions was analysed (1 year) to identify long-term spending in several categories: gas spending, electricity spending, car parking transactions, London emission zone transactions and car insurance transactions. An algorithm was then developed to infer car ownership from these variables. The thresholds for this algorithm were determined by analysing the breakdown of spending across each category *(tables restricted)* and are displayed in Table 2.1.

Assignment	Gas Spend	Electric Spend	Transactions in other car fields
Gas	>£ 500	£0	>= 0
Gas	>£ 100	£0	>= 2
Hybrid	>£ 300	>£100	>= 0
Hybrid	>£ 100	>£0	>= 2
Electric	£ 0	>£100	>= 0
Electric	£ 0	>£0	>= 2

Table 2.1: Table showing algorithm thresholds for car ownership assignment

The emissions due to car ownership were taken as the average carbon cost for an individual's vehicle type (gas, hybrid, electric), divided by its expected lifetime in months, multiplied by the number of months in the study period (6).

The average manufacturing cost of a standard gasoline vehicle is 5.6 tonnes of CO2, a plug-in hybrid is 6.7 tonnes of CO2, and a battery electric vehicle is 8.8 tonnes of CO2 [34]. The lifespan of the average car in 2015 was 13.9 years, recorded by looking at the average car age at the point of being scrapped [35]. This corresponds to a period of 166.8 months, since the study period 6 months, the effect due to car ownership for gas vehicles will be $\frac{5.6 \times 10^6 \times 6}{13.9 \times 12} = 20144g$, the effect due to hybrid car ownership will be $\frac{6.6 \times 10^6 \times 6}{13.9 \times 12} = 23741g$, and the effect due to electric car ownership will be $\frac{8.8\times10^{6}\times6}{13.9\times12} = 31655g$,

It was deemed that 88,812 individuals own a car out of the sample, giving a car ownership proportion of 54.6%. This proportion corresponds perfectly to the actual proportion of car ownership in London, with an estimated 54% of households having access to a car [36]. The breakdown of ownership by car type is in the appendix. There is an issue identifying electric and hybrid (EV) car owners. This issue is potentially the result of many individuals having charging points at home that are not accounted for, so a large amount of plug-in electricity usage is not seen. As a result, many hybrid owners will be mislabelled as gas vehicle owners.

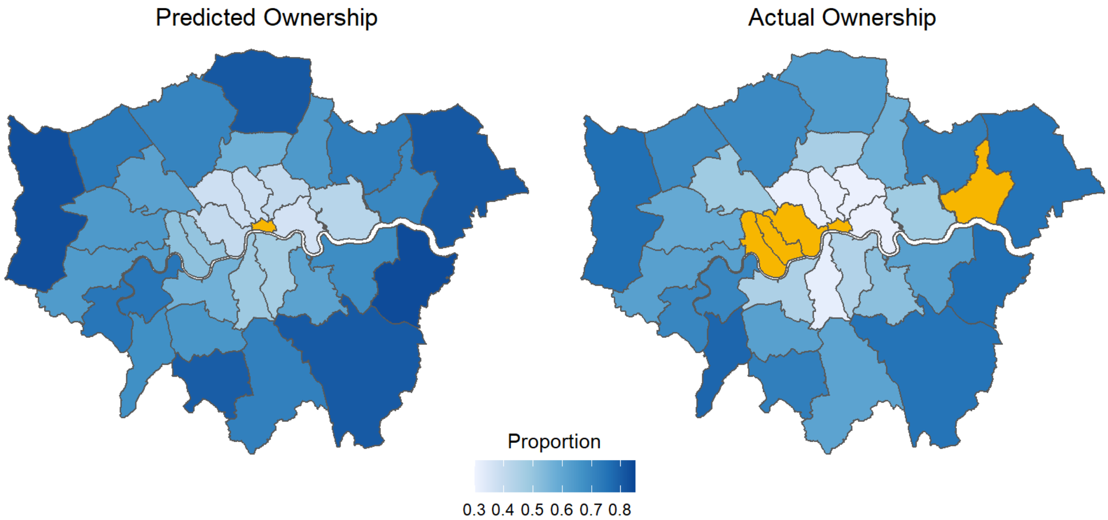


Figure 2.7: Predicted Vs True car ownerships in London

The distribution of predictions shown in Figure 2.7 closely resembles reality. The orange areas are NA values. It can generally be observed that car ownership within Central London is much lower than the ownership seen in the outer regions. Bexley was correctly identified as the LAD with the highest ownership, with predicted ownership of 83% vs actual ownership of 84%. Tower Hamlets was also successfully identified as the LAD with the lowest ownership, with the predicted ownership matching the actual ownership at 34%. A regression was conducted to compare actual vs predicted ownership and found a significant relationship between the predicted and actual results, with a coefficient of 0.988, R-squared of 0.87 and a p-value < 0.001.

2.2.5 Transport For London Transactions

This category represents the emissions from the individual opting to use public transport, namely, those methods within London, including Transport for London (TfL), docklands light railway (DLR), Thames Link, London Overground, and merchants that accept oyster cards.

Since the vast majority of journeys in this category were TfL, the coefficient is based on TfL emissions. A major issue arises with TfL due to the VisaNet infrastructure and payment caps. Due to the high demand at tube stations and on buses, transactions need to be processed quickly, so they are not properly processed at the point of sale. Instead, the transactions are processed at 3 am the following day. This delay means that transactions can be grouped, and a daily cap can be applied to fees [37]. For this reason, the number of journeys and individual journey prices cannot be observed. Therefore, it was deemed that calculating the emissions per pound spent on TfL was the most appropriate method as the daily cap depends on the zones covered; thus, total spending is likely to correlate with distance travelled (as will carbon emissions). In 2020, the total emissions from TfL was $1.04 \times 10^{12} g$ [38], and the total revenue from ticket sales was £4.9b [39], giving an overall impact of 212g/£.

To adapt baselines, a figure of 44g/km of CO2 emissions was considered [40] as provided by TfL in a freedom of information request.

The specific merchants relating to this category were identified and selected.

2.2.6 Taxi

This category represents the emissions from the individual opting to take taxis (including Black Cabs, Ubers, and similar).

The emissions from taxi usage were taken as a function of the number of trips the individual took. The average length of a taxi ride within London was reported to be 5.36km [41]. An assumption was made that the average taxi contained two passengers at any time. With the modal vehicle used as a taxi in London being a Toyota Prius [41], the carbon emissions were deemed 95g/km [42].

The MCC for taxis was not clean. It contains a range of services like Uber, Black Cabs and private chauffeur services, which should all be included; it also includes rental cars, which should not be included. Certain taxi apps, such as FreeNow and Uber, also have e-bike hire options that need to be removed.

Several exclusively e-bike merchants were removed. Then the density of average spending at each merchant was analysed.

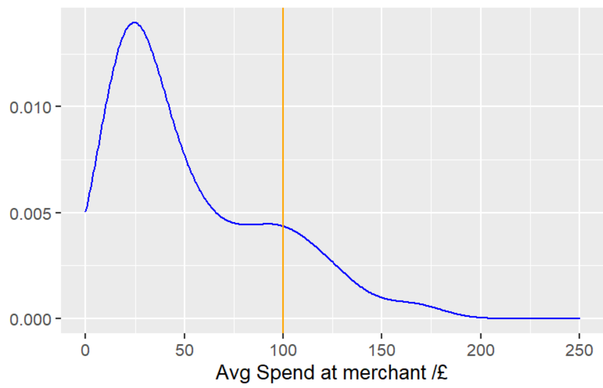


Figure 2.8: Average spend at each merchant in the Taxi MCC

The distribution of average spending at each merchant shows a peak at £25. The curve flattens at £70 - £100 then continues to reduce until it hits 0 around £200. Looking closer at the merchants associated with each level of average spending *(table restricted)* shows that up to £50, the taxis are regular services, those £50 - £100 are generally airport taxi services. Those over £100 are predominantly rental cars or other inappropriate merchants. Thus, a threshold for taxi services was set; only transactions with a merchant from the Taxi MCC that received an average spend of under £100 were considered.

After applying the previous filters, the distribution of spending at various taxi companies was

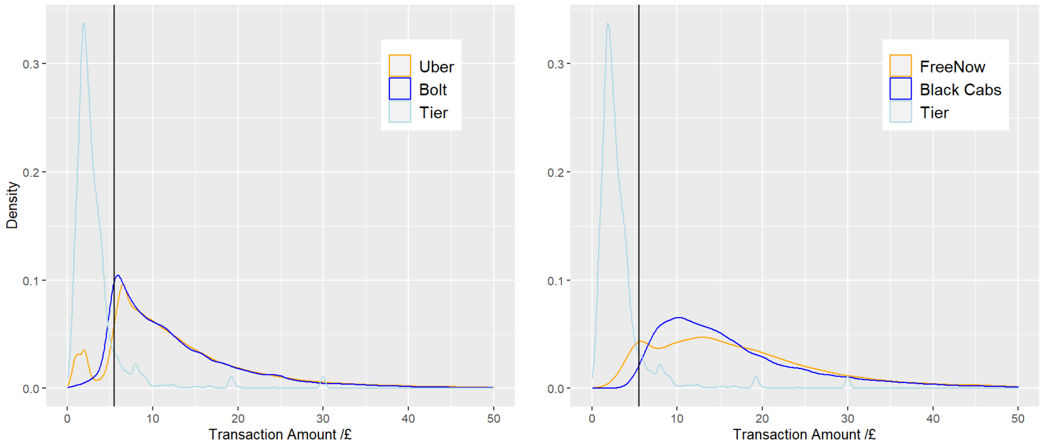


Figure 2.9: Distributions of taxi spends and e-bike spends: Modern and classic payment styles

analysed again to exclude e-bike hire options within specific apps. Due to an array of taxi companies having different payment structures, the distributions of a modern in-app taxi service (Uber) and a classic Black Cab service were analysed. In orange, the distribution of Uber and FreeNow can be seen, services that offer both taxi and e-bike options. In dark blue, Bolt and Black Cabs (via Creative Mobile) can be seen, which are taxi services without e-bike options. Light blue represents e-bike purchases with Tier.

Figure 2.9 shows the two sets of density plots. They are relatively similar. Looking at the left-hand plot, e-bike purchases show a high-density peak at around £2, followed by a sharp decline to low levels at £5; low levels of transactions remain till around £10 where the density approaches 0. Bolt shows a low density up until around £5, which quickly increases to a peak at £6; then, a gradually decreasing function is observed. Uber shows a small local maxima at around £2, decreases, then around £4 appears to increase again to a global maxima at just over £6.

The densities of black cabs and FreeNow transactions (on the right-hand side) showed a similar trend but were less pronounced and shifted to the right by a few pounds. A single filter of transactions equal to or larger than £5.50 was adopted to filter taxi transactions. This number minimises e-bike transactions and maximise taxi transactions. This number is also the minimum fare for uber, which accounts for most of the transactions in this category [43]. The minimum fare for black cabs is slightly lower at £3.80 [44] despite prices appearing to be slightly higher.

This gives an overall direct impact of:

$$n_i \times 5.36 \times (95 - \mathbb{1} \cdot 44) := \begin{cases} 1 & \text{if adjusted for baseline,} \\ 0 & \text{if unadjusted.} \end{cases} \tag{2.3}$$

Where n is the number of transactions an individual makes within the period. This equation gives a total coefficient of 509g per transaction or 273g after accounting for the baseline. It should be noted that behaviours like taxi usage have both a direct effect as a result of their carbon emissions and an indirect effect as a result of causing excess congestion on the roads. For this report, only direct effects are considered.

2.2.7 E-bikes

This category represents those who used communal E-bikes or E-scooters during the study period. Throughout this report, this category will be referred to as e-bikes to avoid confusion (but it also encompasses e-scooters).

The carbon emissions for E-bikes and E-scooters are slightly different, but the behaviour should be considered the same. The median distance travelled per journey is 3.23km for an E-bike and 3.08km for an E-scooter. An E-bike emits an average of 83g/km while an E-scooter emits an average of 106g/km [45]. An assumption was made that an equal amount of e-bikes and e-scooters are used.

$$n_i \times \frac{(83 - \mathbb{1} \cdot 44) \times 3.23 + (106 - \mathbb{1} \cdot 44) \times 3.08}{2} := \begin{cases} 1 & \text{if adjusted for baseline,} \\ 0 & \text{if unadjusted.} \end{cases} \tag{2.4}$$

Where n is the number of transactions an individual makes within the period. This is 297g of CO2, or 158g after accounting for the baseline.

Since only specific merchants have permission to supply communal e-scooters and e-bikes, transactions from these merchants were selected. Additionally, several apps, including Uber and FreeNow, allow for e-bike/e-scooter hire, so these transactions under £5.50 were also selected.

2.2.8 Santander Cycles

The use of London's Santander cycle hire scheme accounts for the emissions due to communal bike services.

The carbon emissions from Santander cycles, excluding those emissions from the consumption of extra food by an individual in order to power said bike is approximately 5g/km [46] and the average trip in 2018 was reported to be 4.2km [47]. These emissions are mainly the result of the fleet of vans servicing and relocating bikes. A Santander cycle allows for unlimited rides within 24 hours, so it will be assumed that the average person takes two journeys within this period.

$$\frac{\pounds P}{2} \times 2 \times (5 - \mathbb{1} \cdot 44) \times 4.2 := \begin{cases} 1 & \text{if adjusted for baseline,} \\ 0 & \text{if unadjusted.} \end{cases} \tag{2.5}$$

Where £P is the amount spent on this behaviour. This gives a coefficient of 42g per pound spent or -328g per pound spent after accounting for the baseline.

Using the amount spent allows those with yearly passes costing £90 [48] to have 45 days worth of cycle hire for the study period inferred, totalling 90 journeys, which seems reasonable.

The single merchant corresponding to TfL Santander Cycles was selected.

2.2.9 Train Transactions

This category represents the emissions from catching trains.

Train tickets vary in price hugely depending on the area and how far in advance they were purchased, so the number of train transactions was taken. The average rail trip emits 79g/km of CO2 [49] and the average journey length is 48.28km [50]. This gives a coefficient of 3814.12g per journey. Since train travel is a baseline, the emissions adjusted for the baseline is 0g.

2.2.10 Coach Transactions

These transactions represent the emissions from those catching coaches.

The average coach trip emits 65.6g/km of CO2 [49], and the average journey length was assumed to be the same distance as the average train journey at 48.28km [50]. This gives a coefficient of 3167.168g of CO2 per journey, or -649g per journey when adjusting for the baseline (train travel).

Visa poorly categorised these transactions, so three specific merchants that comprise most of the transactions were identified and selected: National Express, MegaBus, and FlixBus.

2.2.11 Clothes Purchased

This category represents the emissions due to clothes purchased during the period.

Due to different clothing stores charging different amounts and different items of clothing being set at different price points, the amount spent was deemed a bad indicator. The total number of purchases at clothes stores was taken instead. Each purchase was considered a 'standard basket' of a T-shirt and a pair of jeans totalling a weight of around 1kg, which would deliver a carbon emission of 33700g [51]. These emissions would be similar to a jacket, suit, or a heavy pair of shoes [52].

An MCG existed for clothing. However, several MCCs had to be removed; these include costume hire and wig shops since these serve different purposes and likely have very different environmental impacts. A few miscategorised merchants were removed, and the rest of the MCG was used.

While some level of clothing could be set as a baseline since individuals occasionally need new items, this was deemed unnecessary. Since no other behaviour would be altered, it would impact each score identically, thus changing nothing relatively.

2.2.12 Food Rescue

This category represents the carbon saved from an individual partaking in food 'rescue'. These services encourage individuals to purchase food that would otherwise end up in the bin.

Only two merchants were identified: Too Good To Go and OddBox. Too Good To Go boasts 2495g of carbon saved on average per meal [53] while OddBox boasts between 3000g and 8000g [54]. Since many individuals will make multiple purchases at once, a coefficient to reflect this behaviour was chosen to be 4000g.

2.2.13 Organic Produce

This category considers transactions from organic-only retailers. While shops like Tesco and Waitrose offer organic produce, these transactions cannot be seen. Instead, companies such as Able and Cole and Riverford offer exclusively organic goods.

While organic food production is generally acknowledged to be better for the environment due to the few chemicals used in the agricultural process, this does not immediately translate to lower carbon emissions [55]. For this reason, it is not considered in the environmental, attitudinal score. The behaviour was still mapped from transactions to investigate the ecological correlations.

2.2.14 Locally Purchased Produce

Similarly to organic produce, this category focuses on locally produced food from farm shops, butchers and greengrocers. These merchants were poorly categorised, so this category only includes grocery shops with the word 'greengrocer' or 'butcher' in the name. This categorisation was far from perfect.

Similarly to organic produce, local food consumption generally does not lower an individual's carbon footprint [56] due to more intensive farming methods needed to grow food locally, balancing out the emissions due to food miles. For this reason, it is not considered in the environmental, attitudinal score. The behaviour was still mapped from transactions to investigate the ecological correlations.

2.3 Behaviour Exploration

The behaviours, impacts, and baseline adjusted impacts are in table 2.2.

Behaviour	Coefficent	Adjusted for Baseline
Air Travel	1000g/$	-
Gasoline Purchase	1490.56g/£	887.03g/£
Plug-in Electricity	433.59g/£	1.96g/£
Car ownership: Gas	201439g	-
Car ownership: Hybrid	241007g	-
Car ownership: Electric	315412g	-
TfL	212g/£	**0**
Taxi Transaction	509.2g	273.36g
E-bike Transaction	297g	158.18g
Santander Bike Transaction	42g	-327.6g
Train Transaction	3814.12g	**0**
Coach Transaction	3167.168g	-672.952g
Clothes Transaction	33700	-
Food Rescue Transaction	0	-4000g

Table 2.2: Coefficients of behaviours and baseline adjustments

The score and each behaviour were explored for their correlation with domestic CO2 emissions per capita in tonnes and asthma prevalence. In this report, domestic CO2 emissions are the total CO2 emissions in tonnes produced within each LAD due to fuel usage within the home (heating and electricity) per capita over the last 15 years. A longer time span was taken as the coefficients of the yearly estimates vary year on year. Since no data for 2022 was available, a longer period allows for the general trend in emissions to be captured. That said, there is near perfect correlation between 2019 domestic emissions and the long-term domestic emissions (see appendix 6.22), so this has little impact. A sensitivity analysis was performed using all CO2 emissions produced within the LAD, including sources like manufacturing and transport.

2.4 Scoring

The overall environmental score was the sum of the emissions from the behaviours of interest after accounting for baselines. Three scores were developed; the primary score including all features except air travel. The second score includes air travel, and the third score excludes cars. The equations of these scores may be seen below.

$$\textbf{Score} = \textbf{£P}_{Gas} \times 887.03 + £P_{EV\ electricity} \times 1.96 + \mathbb{1}_{Gas\ Car\ Owned} \times 201439 + \mathbb{1}_{Electric\ Car\ Owned} \times 315412 + \mathbb{1}_{Hybrid\ Car\ Owned} \times 241007 + n_{Taxi} \times 273.36 + n_{e-bike} \times 158.18 + n_{Clothes} \times 33700 - frac£P2 \times 327.6 - n_{Coach} \times 672.952 - n_{Food\ Rescue} \times 4000$$

$$\textbf{Score (with planes)} = \$P_{Air\ Travel} \times 1000 + £P_{Gas} \times 887.03 + £P_{EV\ electricity} \times 1.96 + \mathbb{1}_{Gas\ Car\ Owned} \times 201439 + \mathbb{1}_{Electric\ Car\ Owned} \times 315412 + \mathbb{1}_{Hybrid\ Car\ Owned} \times 241007 + n_{Taxi} \times 273.36 + n_{e-bike} \times 158.18 + n_{Clothes} \times 33700 - frac£P2 \times 327.6 - n_{Coach} \times 672.952 - n_{Food\ Rescue} \times 4000$$

$$\textbf{Score (without cars)} = n_{Taxi} \times 273.36 + n_{e-bike} \times 158.18 + n_{Clothes} \times 33700 - frac£P2 \times 327.6 - n_{Coach} \times 672.952 - n_{Food\ Rescue} \times 4000 \quad (2.6)$$

Where n is the number of transactions, and £P (or $P) is the total spend on a behaviour.

Three scores were necessary for two reasons. Firstly, air miles are unlikely to affect an individual's local area, while their daily activities are; this means there should be a score with and without flights. Then, some individuals rely on vehicles for work, so ownership for some people is a necessity instead of a luxury. As a result, a third score was developed to exclude the use and ownership of vehicles.

The higher the score, the worse the environmental behaviours and the higher the expected carbon emissions.

2.5 Secondary Analysis: Asthma Prevalence

Asthma prevalence is heavily linked to pollution and the environment [57]. For this reason, the scores will also be assessed for correlation with this outcome. The expectation is that the higher the score, the higher the prevalence of asthma. The asthma prevalence given is the prevalence percentage (prevalence per 100 people). It is calculated by looking at GP surgeries within each LAD. As a result, some inaccuracy will arise from many individuals not seeing GPs within their LAD.

Results

All maps produced in this section have a corresponding table in the appendix of exact values. Orange on a map represents an NA value. Additionally, regression tables, including coefficients, r-squared and p-values, are in the appendix for all regressions performed.

3.1 Environmental Behavioural Score

Once the scores were made, they were all linearly transformed to be between 0 and 1. A score of 0 represents the most environmentally conscious behaviours, and a score of 1 represents the most negligent environmental behaviours. The three scores all present similar findings.

3.1.1 Score and CO2

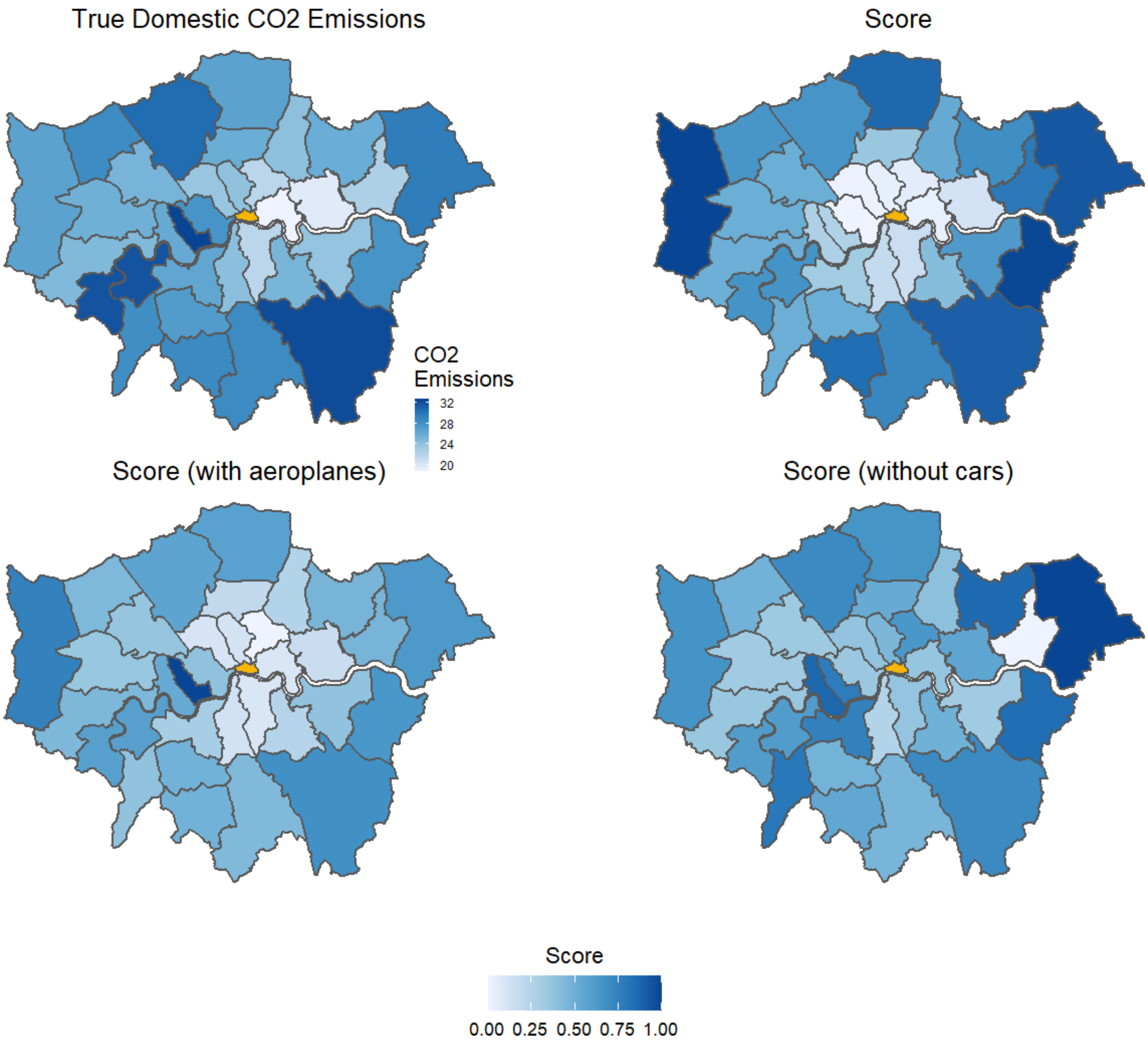


Figure 3.10: Scores vs Domestic CO2 emissions

Figure 3,1 show the distribution of domestic CO2 emissions in London shows that east of Central London has the lowest emissions, with Tower Hamlets at 18.69 per person and Newham at 19.56 per person. Central London also has relatively low emissions, except for West Central London.

The borough of Kensington and Chelsea is the highest emitter, with 32.59 per person. Generally, the larger emitters are in Outer London, with Bromley emitting 32.36 per person and Richmond emitting 31.83 per person.

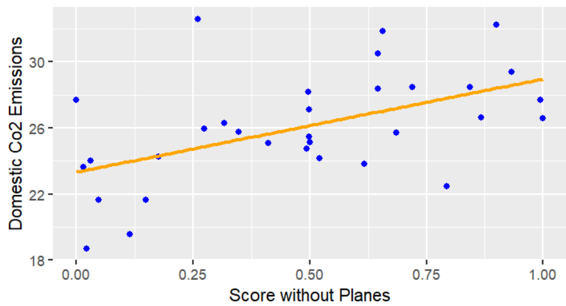
The score (top right of Figure 3.1) by LAD shows a clear pattern. Central London has smaller scores, as reflected by Westminster scoring 0, and Camden, Tower Hamlets, Islington and Hackney all scoring below 0.05. Just out of central London shows a ring of medium scoring LADs, then in Outer London, a couple of LADs receive high scores, including Hillingdon with a score of 1, as well as Bexley, Havering and Bromley, all scoring over 0.9.

The score without aeroplanes shows a similar pattern. While it correctly assigns the highest score to Kensington and Chelsea, scoring 1, it fails to predict the magnitude of emissions in the outskirts of London, with the second highest score going to Hillingdon with a score of 0.75, followed by Bromley and Bexley with scores of 0.67 and 0.64 respectively.

The score without cars has the weakest correlation with the True Domestic CO2 emissions. The score in outer London, excluding eastwards, seems to correlate quite well with actual domestic CO2 emissions. However, East London shows an issue with the magnitude of the scores. In Barking and Dagenham, the score without cars is 0. The next lowest scorer is Lambeth with 0.25, then Ealing with 0.32 making barking and Dagenham notably the lowest scorer. In the true CO2 emissions, Barking and Dagenham is the fifth smallest emitter, with a score of 22.44.

To quantify these correlations, a simple linear regression was run for each score against the outcome of domestic CO2 emissions. The assumptions were checked and were not violated. The results are shown in figure 3.11.

Score	Coef	R-Squared	P-Value
Score	0.049	0.28	0.002
Score (with aeroplanes)	0.047	0.56	< 0.001
Score (without cars)	0.029	0.23	0.006



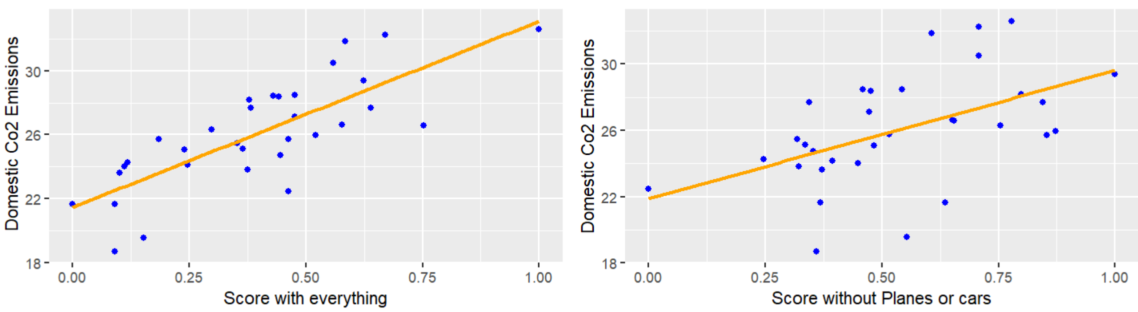


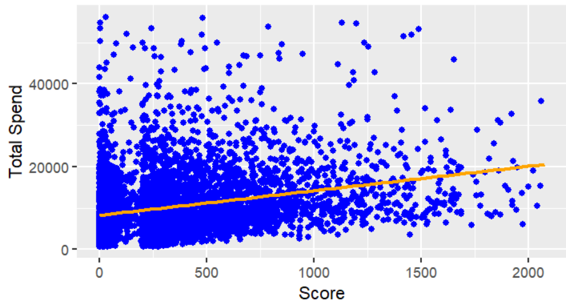
Figure 3.11: Scores vs actual CO2 emissions correlations

Each of the scores showed a significant correlation with domestic CO2 emissions. The score with aeroplanes shows the strongest correlation with an R-squared of 0.56 and a p-value <0.001. The score and the score without cars have notably smaller r-squared of 0.28 and 0.23, respectively, implying that the scores capture less variation in the data. However, the p-values are still significant at 0.002 and 0.006, respectively. The positive coefficients show that the higher the score, the larger the expected CO2 emissions are for a particular LAD.

Figure 3.12 explores how the scores correlate with total spending at the individual level. A positive correlation can be observed, as in previous research by Visa. The plots show 5000 random individuals' scores to better visualise the relationship with their total spending. The plots show heteroskedasticity, which violates the Gauss-Markov assumptions, implying that the given P-value might be artificially small. However, a relationship still appears to exist. At the LAD level, however, no significant relationship exists between score and total spending (see appendix 6.23).

The relationship between total spending and score at the individual level might result from total spending being a major factor in CO2 emissions. Very few people in the UK have no disposable income and live off nothing but essentials. After applying filters for 'active cards', these individuals might have been removed as they will transact less. This means that most individuals captured have an income to make consumer choices with, and these choices almost always increase emissions.

Score	Coef	R-Squared	P-Value
Score	0.0140	0.08	< 0.001
Score (with aeorplanes)	0.0270	0.13	< 0.001
Score (without aeroplanes)	0.0012	0.06	< 0.001



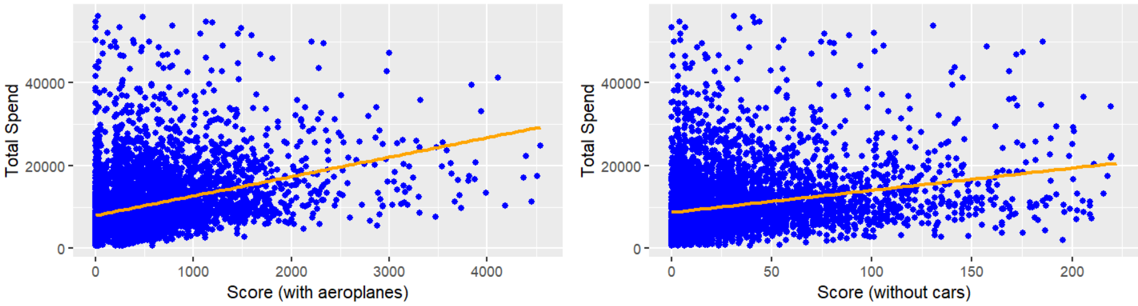


Figure 3.12: Scores vs total spend at the PAN level

A sensitivity test was conducted to investigate the scores against total CO2 emissions within each LAD, including those from industrial processes, transport and tourism. Since most emissions are not related to the individuals residing in the LAD, any correlations observed will strengthen the findings.

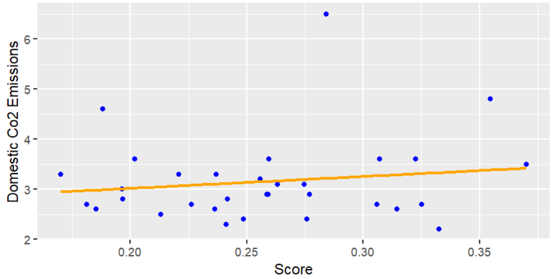
Figure 3.13 shows that both the score and score without cars have no correlation with the total CO2 emissions with p-values > 0.05. However, the score with aeroplanes has a significant correlation, with an R-squared of 0.13 and a P-value of 0.042. The positive correlation suggests that the higher an LAD's score (with aeroplanes), the higher the expected total emissions.

3.1.2 Score and Asthma

Figure 3.14 looks at the distribution of asthma prevalence across London. It shows that West Central London has the lowest rates, with Westminster having a prevalence of 3.36 and Kensington and Chelsea having a prevalence of 3.38; this is possibly the result of having high numbers of private GPs. The LAD with the most considerable prevalence is Sutton with 5.94, followed by Lewisham, Bromley and Harrow with prevalences above 5.30.

Looking at how the environmental scores correlate with asthma, it can be observed that the score with aeroplanes and the score without air travel are highly significant, with p-values of < 0.001 (appendix table 6.6). The score with air travel has a higher r-squared of 0.64 than that without

Score	Coef	R-Squared	P-Value
Score	0.0091	0.02	0.42
Score (with aeroplanes)	0.1100	0.13	0.042
Score (without cars)	-0.0330	0.01	0.59



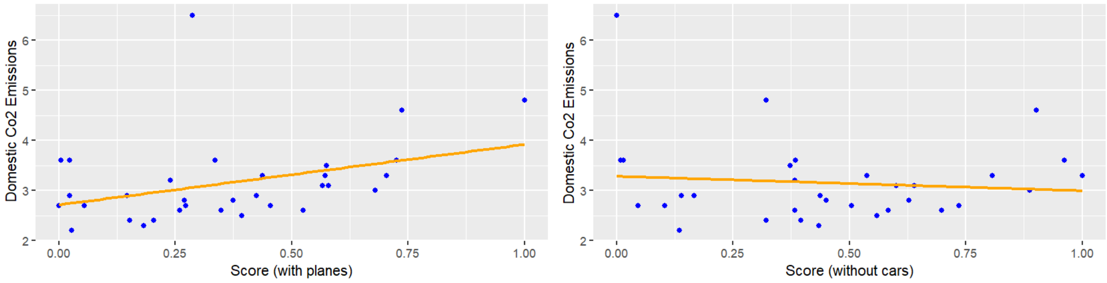


Figure 3.13: Sensitivity Analysis: Scores against total CO2 emissions

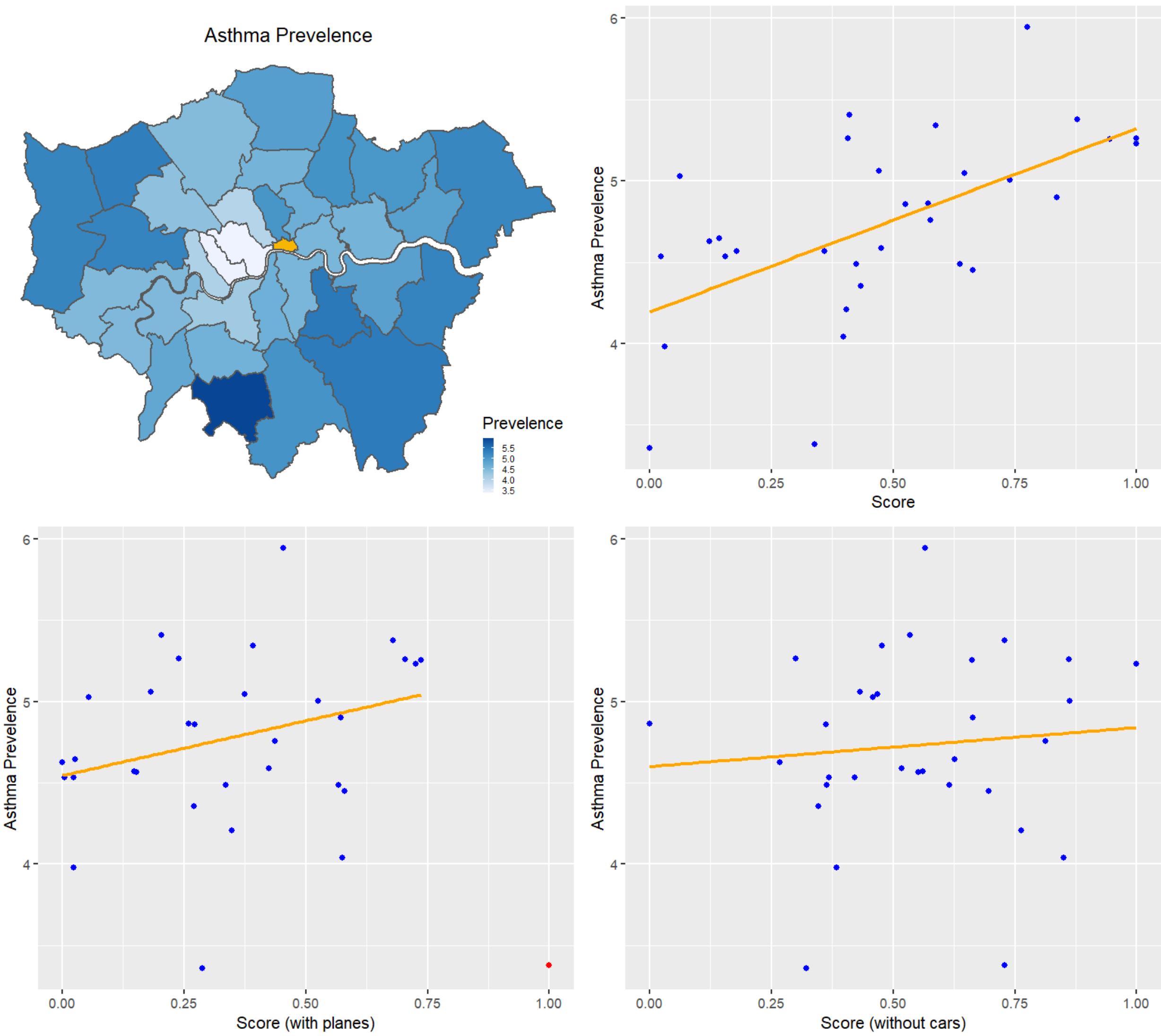


Figure 3.14: Scores vs Asthma Prevalence

air travel of 0.34. The score without air travel or cars performs the worst but is still significant to the 5% level with an R-squared of 0.24. The coefficients are all positive, meaning that keeping everything else constant, the higher the score, the higher the expected asthma prevalence in an LAD.

3.2 Biggest Factors

Figure 3.15 investigates what drives these scores at the LAD level. The plot shows the unscaled CO2 contributions after accounting for baselines (see appendix for unadjusted). It shows that the scores are almost entirely attributable to Clothes purchased, Flights, Gasoline Car Ownership, and Driving. As previously seen on the map, the LAD with the highest emissions is Kensington and Chelsea, with just under 1,200,000g of CO2 per person.

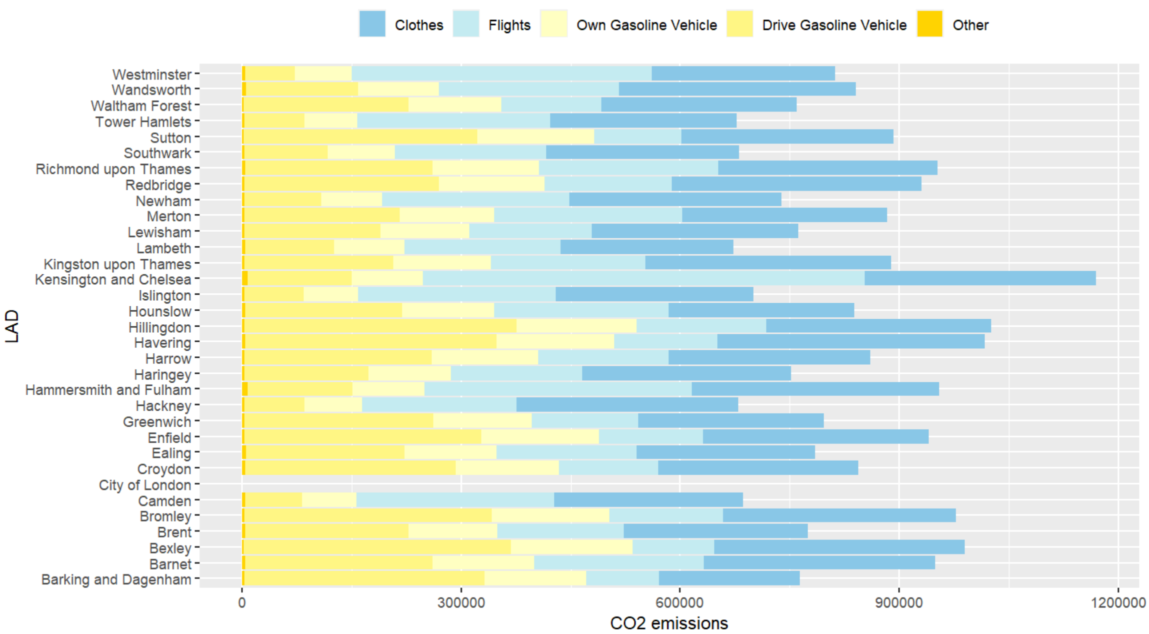


Figure 3.15: Mean score contributions by LAD

Certain behaviours, such as electric car ownership, have a large coefficient at the PAN level but do not drive scores at the LAD level due to the low prevalence of the behaviour. To further understand the results, some behaviours will be investigated further.

3.2.1 Flying

Figure 3.16 shows the distribution of average emissions per person due to flying in each LAD. Additionally, it shows the proportion of individuals who flew by LAD. It can be observed that those in Central London tend to emit slightly more from flying compared to those who live in Outer London. This difference is further reflected in the proportion of individuals predicted to have flown. Kensington and Chelsea has a large coefficient for flight emissions, which average 605kg per person. The second and third largest coefficients for flight emissions are Westminster and Hammersmith, with coefficients of 411kg and 366kg per person, respectively. All other flight emissions at the LAD level are below 300kg. Interestingly, Kensington and Chelsea only have the second highest proportion of individuals who have flown, with 35% compared to Hammersmiths

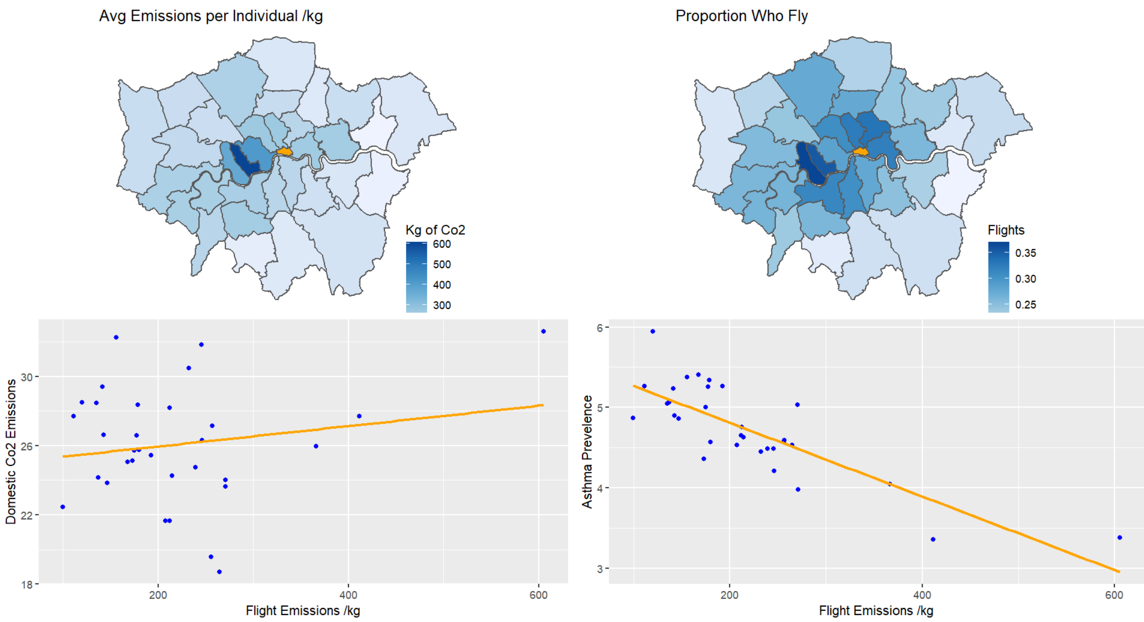


Figure 3.16: Flights by LAD and CO2 correlations

37% despite far higher emissions. This contrast potentially results from the stark wealth differences within Kensington and Chelsea. The borough contains large areas that constitute the top 10% most deprived in the UK and others that constitute the top 10% least deprived in the UK [58], meaning a large group may not fly, while a large group fly lots. Alternatively, those within Kensington and Chelsea might buy more expensive plane tickets, including more first-class tickets or longer-haul flights.

The trend between Flight emissions and Domestic CO2 Emissions shows no significant correlation with a coefficient of 0.0059, R-squared of 0.03 and with a p-value of 0.50. This p-value remains not significant if Kensington and Chelsea is removed (the point on the top right-hand corner). These trends highlight the difference between the scores with and without planes, specifically why Kensington and Chelsea goes from high scoring (but not the highest) to the highest scoring by a large margin when flights are considered.

The trend between flights and asthma shows a strong negative correlation of -0.0046 with an R-squared of 0.65 and a p-value < 0.001. The three aforementioned LADs with the largest emissions give the distribution of results a long right-hand tale.

Further investigation was conducted on the relationship between flight emissions and asthma. The proportion of total spending on 'daily purchases' was added as a feature in the regression (transport and groceries) since this is the approach taken by the VCA team to account for wealth. However, the r-squared remained virtually unchanged, and the coefficient for wealth was not significant, implying wealth probably does not impact this relationship. Another hypothesis is that those with asthma are less able to travel due to their condition. More research should be conducted on the link between air travel and asthma.

3.2.2 Inner London Transport

The majority of individuals took a variety of transportation methods. Within the sample, the most common combinations were: 17434 drove and used TfL, 16024 drove, took TfL and used taxis, 10955 exclusively used TfL, and 10426 individuals only used TfL and Taxis. In total, 80% of the sample used TfL, 63% drove, 51% used taxis, 15% used e-bikes and 2% used Santander bikes.

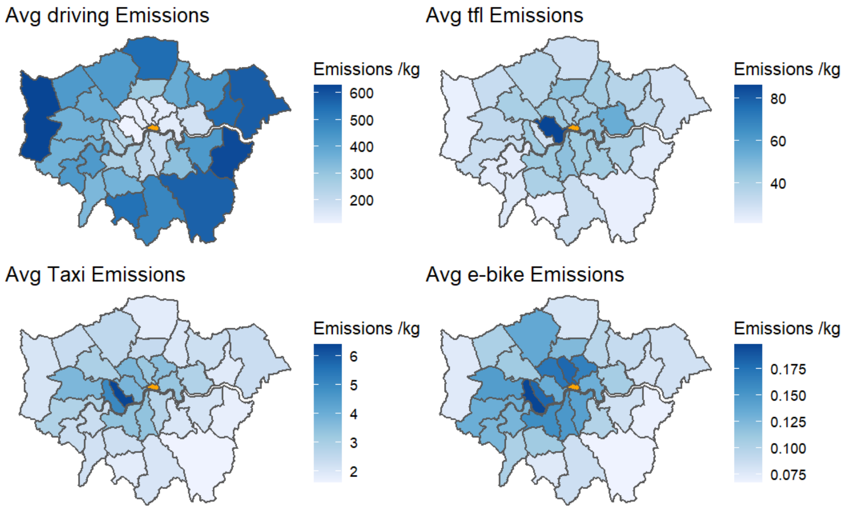


Figure 3.17: Emissions from transport within London by LAD

Figure 3.17 shows a breakdown of the public transport usage of those living within London.

Taxi, TfL, and e-bike usage are all primarily used by those in Central London. Taxis and e-bikes have a slight westward tendency, possibly due to West London being wealthier [59]. TfL has a slight northern tendency. Driving shows the opposite, with central London driving much less than outer London. This disparity in driving could be down to worse public transport links in Outer London and the lack of a Low Emission Zone charge, Ultra Low Emission Zone charge, or Congestion Zone charge [60], which are upheld in Central London.

Kensington and Chelsea have interesting tendencies with high taxi emissions and exceptionally low TfL emissions. This supports the theory that their distribution is heavily affected by wealth.

Driving

Driving was flagged as a highly impacting behaviour. The average score for gasoline-based vehicle owners, electric or hybrid vehicle owners and non-vehicle owners was compared, allowing insight into the consequence of PAN level behaviours.

Figure 3.18 shows a breakdown of carbon emissions at the car ownership level. Those with electric cars emit the most, closely followed by those with gas cars. Emitting much less on average is those with no cars. Some similarities and differences can be observed. Looking at Non-vehicle owners, it can be observed that they have the most emissions from TfL, which logically follows. Those who drive gas cars emit much more through driving their vehicle than those with EV cars. On average, non-vehicle owners emit 194kg of CO2 through flying, similar to those driving gas vehicles with 216kg. On average, those who drive EV vehicles emit much more from flying with 435kg. Electric vehicle owners emit the most, mainly due to their flying habits; this is likely as electric vehicle owners are better off. The similarities between gasoline vehicle owners and non-vehicle owners suggest that similar individuals make up these categories.

Gas cars are much worse than electric cars at the street level due to non-carbon emissions such as nitrogen oxide [61]. However, as previously mentioned, this study focuses exclusively on carbon emissions.

The top two graphs of Figure 3.19 show strong positive correlations between gas spending and CO2, and gas spending and asthma at the LAD level, with an r-squared of 0.24 and 0.4 along with significant p-values of 0.004 and < 0.001, respectively. The bottom row of the graphs shows car

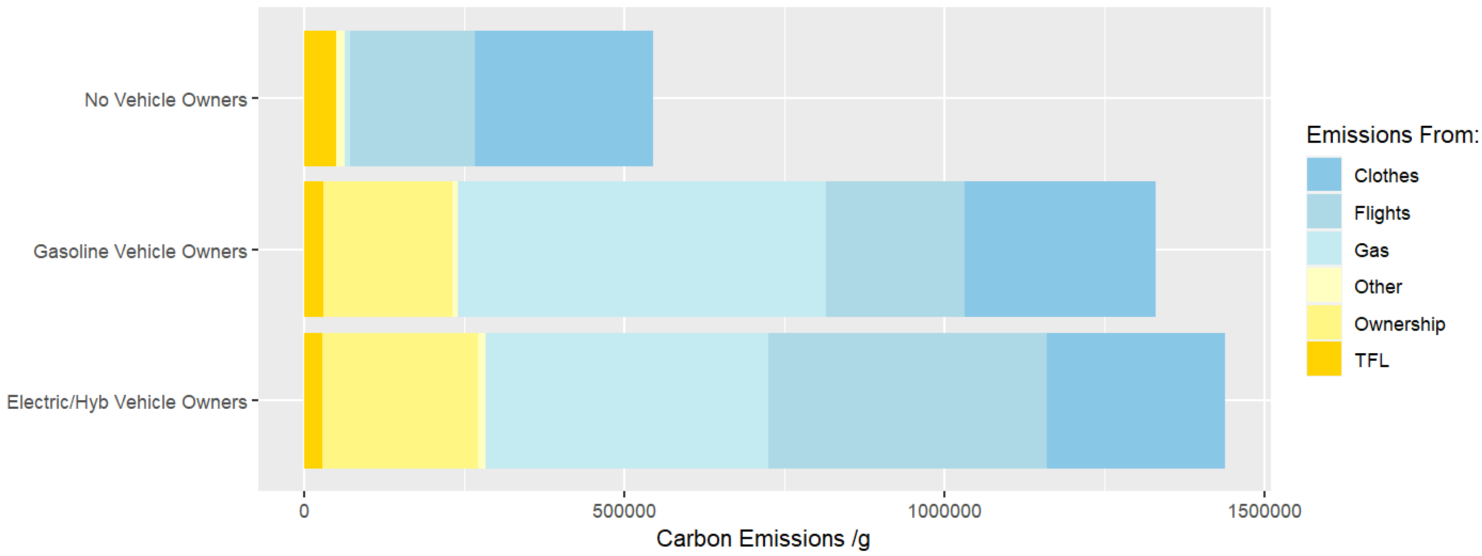


Figure 3.18: Vehicle owners emission comparison

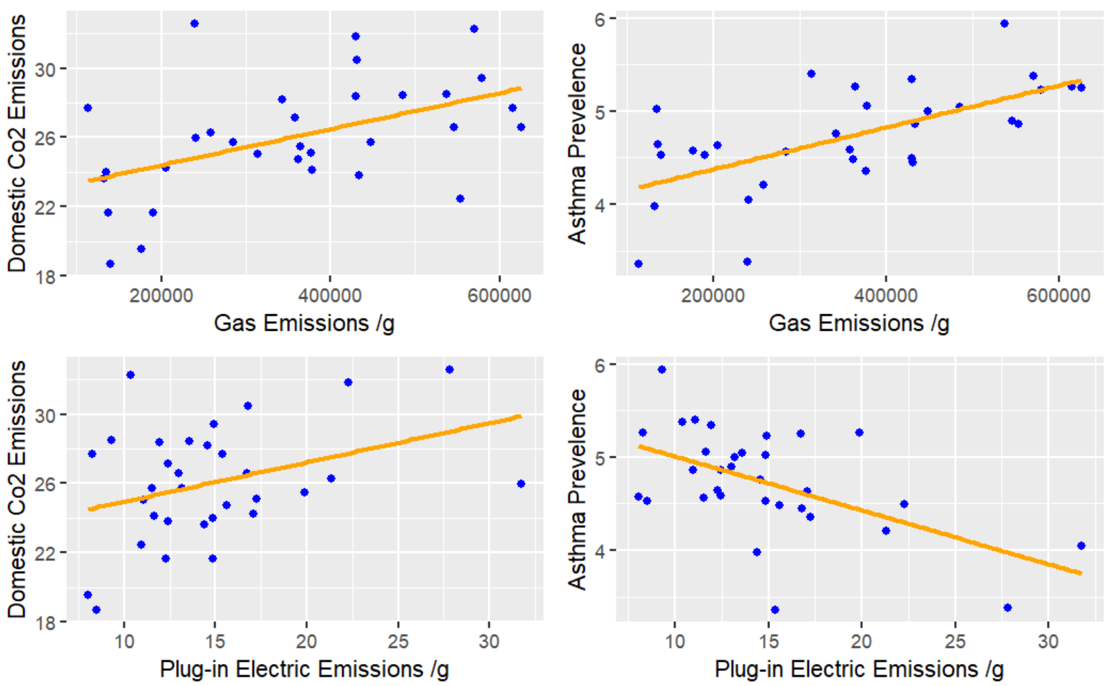


Figure 3.19: Fuel emissions and their correlations with CO2 and asthma prevalence

electricity emissions vs CO2 and Asthma; Electricity purchases positively correlate with CO2 with an r-squared of 0.13 and p-value of 0.045, while electricity has a negative association with asthma with an r-squared of 0.3 and p-value of 0.001 (see appendix table 6.11)

These results mean that for two identical LADs, one with a higher gas purchase keeping everything else constant, one would expect the LAD with a higher gas purchase to have higher asthma rates and higher CO2 emissions. Keeping everything else constant, a LAD with higher electric spending than another is expected to have higher CO2 emissions and lower asthma rates. The correlation between electric cars and asthma could be wealth related, with those who are richer more likely to own an electric car and less likely to suffer from asthma symptoms.

3.2.3 Consumer Behaviour

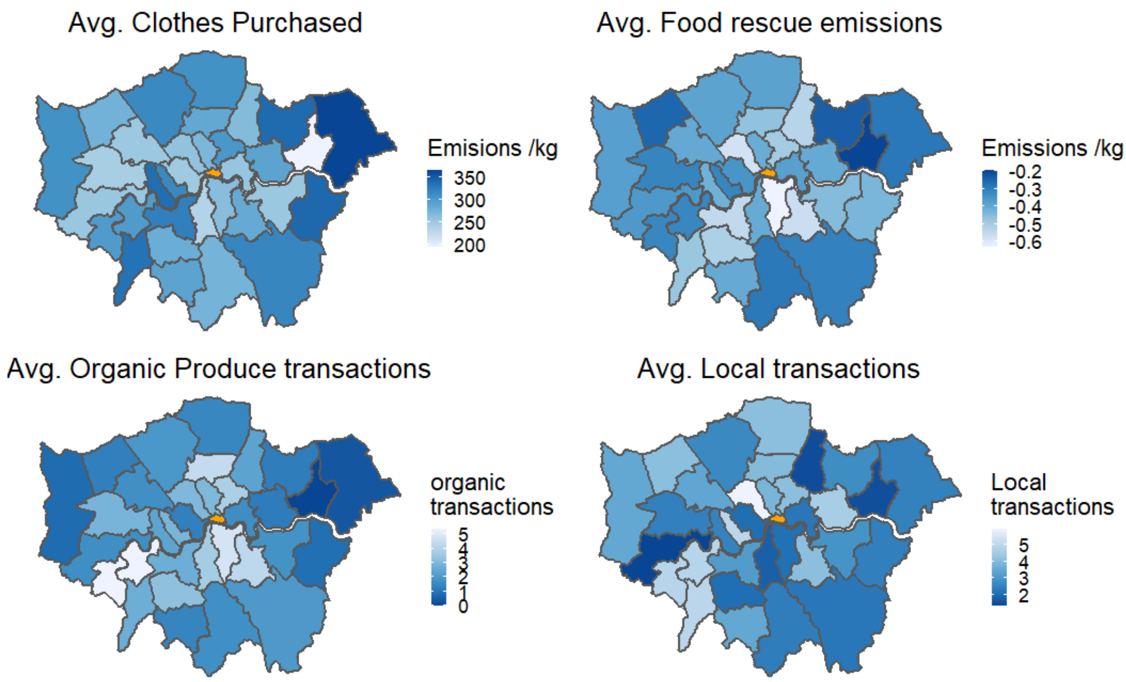


Figure 3.20: distribution of consumer purchases

The maps showing consumer behaviour in Figure 3.2 have some similarities and differences. They all show a relatively even spread of magnitude across London, with no large differences in emissions or transactions between central and outer London, nor do they have any cardinal directional trends. Many LADs are consistent across all plots, and some show some discrepancies. For example, Brent always has a middling magnitude, while Richmond has a near zero magnitude for food rescue emissions but the most organic transactions.

Looking at the correlation matrix between these variables in Figure 3.21, a positive correlation exists between food rescue emissions saved, organic transactions and local transactions. This makes sense since many individuals will be conscious of the foods they eat and perform multiple of these behaviours. In contrast, little to no correlation exists between groceries and average clothes purchased, but clothes purchasing is a very different consumer behaviour, so this is expected.

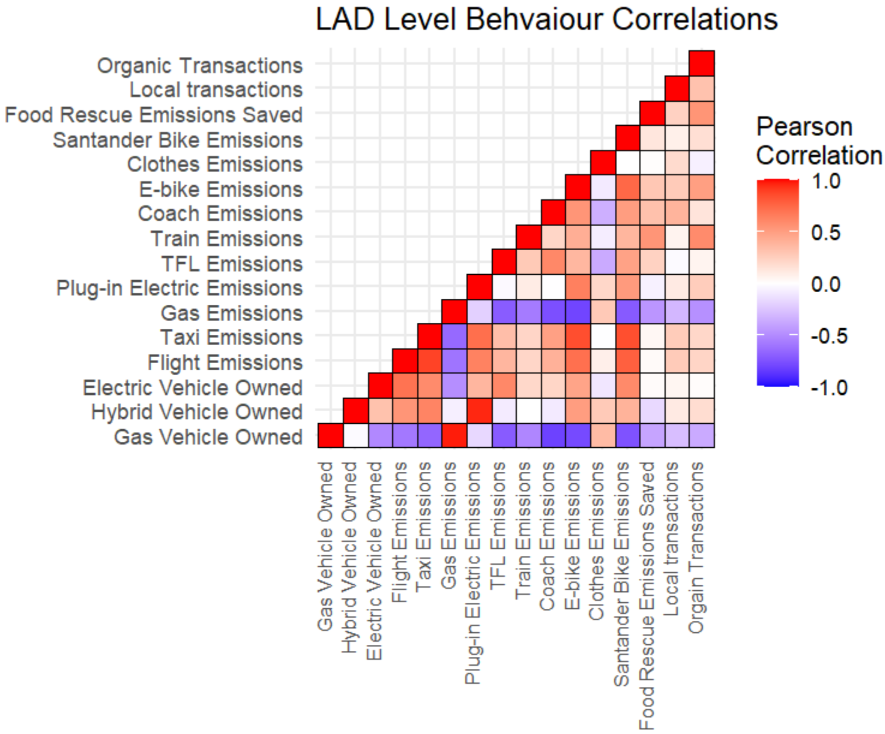


Figure 3.21: correlation between behaviours (LAD level)

3.3 Overlapping behaviours

Due to the sparse nature of the data, correlations between variables are more substantial at the LAD level, as shown in Figure 3.21. PAN-level correlations may be seen in the appendix under Figure 6.25. The first observation is ownership, and the use of gas vehicles negatively correlates with other variables; Using a car means less need to travel by other forms of transport, which would drive this negative correlation. The negative association with food purchases might be confounded by the area lived. Those who live in central London, where the wealth is more heavily concentrated, are more likely to shop organically or locally but are less likely to own a car due to better transport links. Clothes purchases are the only behaviour positively associated with gas vehicle ownership and use; this is likely a result of a car facilitating more shopping.

The majority of other variables hold positive correlations with each other, demonstrating the impact of total spending. The more a person performs one activity, the more likely they are to perform another behaviour too. This belief is confirmed by the PAN level correlations (appendix), where many variables hold (weak) positive correlations.

Discussion

4.1 Transactional Data and Environmental Behaviour

Revising the initial aims: Can transactional data be used to construct an environmental and behavioural score at the PAN and LAD levels? The results of this study were clear. Environmental behaviours can be determined from transactional data, and they can successfully be used to construct a score which reflects the environmental attitude and the PAN and LAD levels.

Do ecological correlations exist between environmental behaviours and asthma prevalence? Yes, a significant positive correlation exists between environmental behaviour and asthma prevalence.

Looking at how these scores correlate with total spending on the LAD level show no significant relationship. However, a significant relationship between average spending and score can be seen at the PAN level. Keeping everything else constant, seeing two individuals with different scores, the one with the higher score has a higher expected total spend.

This links to previous research that suggested that the wealthier an individual, the higher their carbon emissions [62]. On the extreme side, a 2020 report from Oxfam found that the wealthiest 10% of individuals were responsible for approximately half of global emissions [63]. Within this study, the impacts of wealth can be observed. Repeatedly, Kensington and Chelsea have high spending on specific categories associated with wealth, namely flight emissions and taxi emissions. The average total spend within Kensington and Chelsea of £16413.04, around £2000 more than the next highest LAD, while the sample average is £10649.49.

4.2 Limitations

4.2.1 Data Collection

While Visa collects vast amounts of data, there are some major limitations to the study, primarily surrounding data collection. Transactional data provides information on the merchant; this is useful when the merchant serves a niche purpose with limited products, but when they sell multiple different things, there is no way of knowing what an individual purchase is for. This becomes an issue at places such as fuel stations, where if a purchase is made inside, it is impossible to differentiate between a fuel purchase and someone buying dinner.

Furthermore, transactions that occurred through other mediums, such as Cash or MasterCard, cannot be seen. Throughout the analysis, it was assumed that the transactions observed on an individual's Visa account represent all their purchases across all media. The assumption was also made that the 'active' cards represent all individuals when in reality, the sample is likely to be unrepresentative of the less wealthy. These issues are not easily fixable.

Only a small subset of behaviours could be calculated. In 2020, just 27% of carbon emissions were as a result of transport [64], which in this report made up over 90% of the score. It would be helpful to further understand individuals' diets, among other things, but this is impossible with transactional data.

4.2.2 Data Processing

There are further issues with the data from the data processing side. This analysis relied on purchases being distributed into accurate MCCs and MCGs. While, on the whole, they are reasonably good due to having a data science team dedicated to their allocation, they are far from perfect. Many of the behaviours mapped were based upon this prior categorisation by Visa; if a large merchant was mislabelled and put into the wrong category, it could potentially lead to drastically different results and the score misrepresenting reality. While an effort was made to remove inappropriate merchants, they could not all be screened individually.

4.2.3 Transaction Data Bias

The European Central Bank conducted a study that found that preference for electronic payment technology over other means lowers with the transaction's size, educational level, age and income of the individual in question. [65]. The implication is that the selected individuals are likely younger, more educated and more wealthy than the average individual.

4.2.4 Analysis

The analysis was the source of several limitations. First and foremost, many significant correlations were computed. Nevertheless, this does not imply causation. Most behaviours have had causation established in other research. Secondly, an ecological analysis was used to validate the scores. Ecological studies have the problem that they cannot directly attribute coefficients to individuals (ecological fallacy); thus, the reliability of the scores, especially at the PAN level, is limited.

4.2.5 Implications

Allowing individuals to see their score could impact their behaviours, potentially helping them make decisions to reduce their carbon footprint. Individuals could see their scores through the launch of a Visa app, where users can link one or multiple cards. This could give users information on their behaviours and how much carbon they would save by performing alternative actions. Rewards and incentives could be offered to help individuals follow these behaviours. For example, an individual with high spending on taxis could be offered cashback on e-bike rides. This approach would allow the individual to optionally enter in more personal details, which will allow for more personalised suggestions (e.g. not suggesting the elderly use e-bikes) and more detailed analytics of behaviours.

4.3 Project Extensions

There are extensions to this project that could strengthen the findings and help provide a more in-depth understanding of environmental behaviours in the UK. Extending the scoring to the rest of the UK would be insightful. However, it would require mapping out more region-specific behaviours, with an extra emphasis on methods of transport since this study has indicated that they are critical. Another extension would be to see how these findings compare to 2019 or other previous years; it would be hoped that the LAD breakdown would be somewhat proportional, even if the magnitude of the scores were different. A comparison would allow for understanding the difference in behaviours between now and pre-COVID-19 pandemic, which would help answer questions like: Are consumers becoming more conscious of the impact of their decisions?

More sensitivity analyses would have helped strengthen the findings. Sensitivity analyses that could have been performed include looking at the correlation with other metrics, such as domestic water usage. Water usage is an environmental outcome independent of carbon emissions. Another sensitivity test could have altered the scoring and conducted the analysis with extreme values included in their original form, or with extreme values removed completely (instead of being shrunk to the 99th percentile).

Conclusion

In this study, it has been shown that transactional data can be used to map individuals' behaviours and determine their impact on the environment. This impact can be reflected in a numeric score. This score has been validated on the LAD level by establishing a significant positive correlation with each LAD's actual domestic CO2 emissions. This score is independent of total spending at the LAD level, while at the PAN level, a relationship exists with total spending.

It has also shown that at the LAD level, the score is positively associated with asthma prevalence, meaning that the higher the score (meaning, the worse the behaviours), the higher the expected prevalence of asthma.

There were four main drivers behind the score at the LAD level. These were clothes purchases, air travel, gasoline vehicle ownership and gasoline vehicle use. When looking at raw emissions at the LAD level, TfL usage also contributes a substantial amount of emissions. These high contributions reflect simultaneous high emissions and uptake within the population.

At the PAN level, those who own vehicles have substantially more emissions on average than those who do not. Those who drive EVs emit the most, largely due to a vast increase in flight emissions from the EV owners; this relationship is likely confounded by wealth.

The findings' strengths are limited due to a lack of cardholder information and due to being validated at the ecological level.

Appendix

Area Code	Area Name	Population	New Classification	Original
E09000001	City of London	0.00098	0.0011	0.023
E09000002	Barking and Dagenham	0.025	0.0083	0.0052
E09000003	Barnet	0.044	0.049	0.029
E09000004	Bexley	0.028	0.031	0.021
E09000005	Brent	0.039	0.028	0.013
E09000006	Bromley	0.038	0.042	0.043
E09000007	Camden	0.024	0.026	0.025
E09000008	Croydon	0.044	0.049	0.03
E09000009	Ealing	0.042	0.036	0.021
E09000010	Enfield	0.038	0.039	0.022
E09000011	Greenwich	0.033	0.036	0.02
E09000012	Hackney	0.029	0.032	0.017
E09000013	Hammersmith and Fulham	0.021	0.023	0.016
E09000014	Haringey	0.03	0.033	0.019
E09000015	Harrow	0.03	0.023	0.012
E09000016	Havering	0.03	0.032	0.021
E09000017	Hillingdon	0.035	0.037	0.03
E09000018	Hounslow	0.033	0.027	0.017
E09000019	Islington	0.025	0.027	0.02
E09000020	Kensington and Chelsea	0.016	0.018	0.014
E09000021	Kingston upon Thames	0.019	0.021	0.023
E09000022	Lambeth	0.036	0.04	0.025
E09000023	Lewisham	0.034	0.038	0.021
E09000024	Merton	0.024	0.027	0.018
E09000025	Newham	0.04	0.027	0.014
E09000026	Redbridge	0.035	0.032	0.017
E09000027	Richmond upon Thames	0.022	0.025	0.025
E09000028	Southwark	0.035	0.039	0.02
E09000029	Sutton	0.024	0.026	0.023
E09000030	Tower Hamlets	0.035	0.029	0.016
E09000031	Waltham Forest	0.032	0.031	0.015
E09000032	Wandsworth	0.037	0.041	0.033
E09000033	Westminster	0.023	0.026	0.33

Table 6.3: Table showing the proportions assigned to each LAD by both methods and the actual population proportions.

LAD code	LAD	Actual Ownership Proportion	Predicted Ownership Proportion	Predicted Gasoline vehicle ownership	Predicted hybrid ownership proportion	Predicted electric ownership proportion
E09000001	City of London	NA	NA	NA	NA	NA
E09000002	Barking and Dagenham	NA	0.7	0.69	0.0082	0
E09000003	Barnet	0.69	0.7	0.69	0.013	0
E09000004	Bexley	0.76	0.84	0.83	0.0065	0
E09000005	Brent	0.48	0.62	0.6	0.012	0
E09000006	Bromley	0.75	0.81	0.8	0.0094	0
E09000007	Camden	**0.31**	0.38	0.37	0.0092	0.00042
E09000008	Croydon	0.61	0.72	0.71	0.012	0.00024
E09000009	Ealing	0.59	0.64	0.63	0.015	0
E09000010	Enfield	0.64	0.81	0.8	0.01	0
E09000011	Greenwich	**0.62**	0.68	0.67	0.011	0
E09000012	Hackney	0.31	0.4	0.39	0.0088	0
E09000013	Hammersmith and Fulham	NA	0.51	0.49	0.021	0.00048
E09000014	Haringey	**0.46**	0.57	0.56	0.0098	0
E09000015	Harrow	0.69	0.74	0.73	0.011	0
E09000016	Havering	0.75	0.81	0.8	0.012	0
E09000017	Hillingdon	0.76	0.83	0.82	0.012	0
E09000018	Hounslow	0.62	0.64	0.63	0.012	0
E09000019	Islington	0.31	0.38	0.37	0.009	0.00041
E09000020	Kensington and Chelsea	NA	0.5	0.48	0.019	0.00062
E09000021	Kingston upon Thames	0.78	0.67	0.66	0.011	0
E09000022	Lambeth	0.32	0.49	0.48	0.01	0
E09000023	Lewisham	0.52	0.61	0.6	0.0087	0
E09000024	Merton	0.62	0.65	0.65	0.0078	0
E09000025	Newham	0.48	0.43	0.42	0.0074	0
E09000026	Redbridge	0.72	0.72	0.71	0.01	0
E09000027	Richmond upon Thames	0.7	0.74	0.73	0.015	0
E09000028	Southwark	0.44	0.47	0.46	0.011	0.00034
E09000029	Sutton	0.72	0.8	0.79	0.0077	0
E09000030	Tower Hamlets	0.31	0.36	0.36	0.0058	0
E09000031	Waltham Forest	0.57	0.64	0.64	0.0078	0
E09000032	Wandsworth	0.45	0.57	0.55	0.015	0
E09000033	Westminster	NA	0.4	0.39	0.011	0.0012

Table 6.4: Proportion of individuals owning cars by LAD

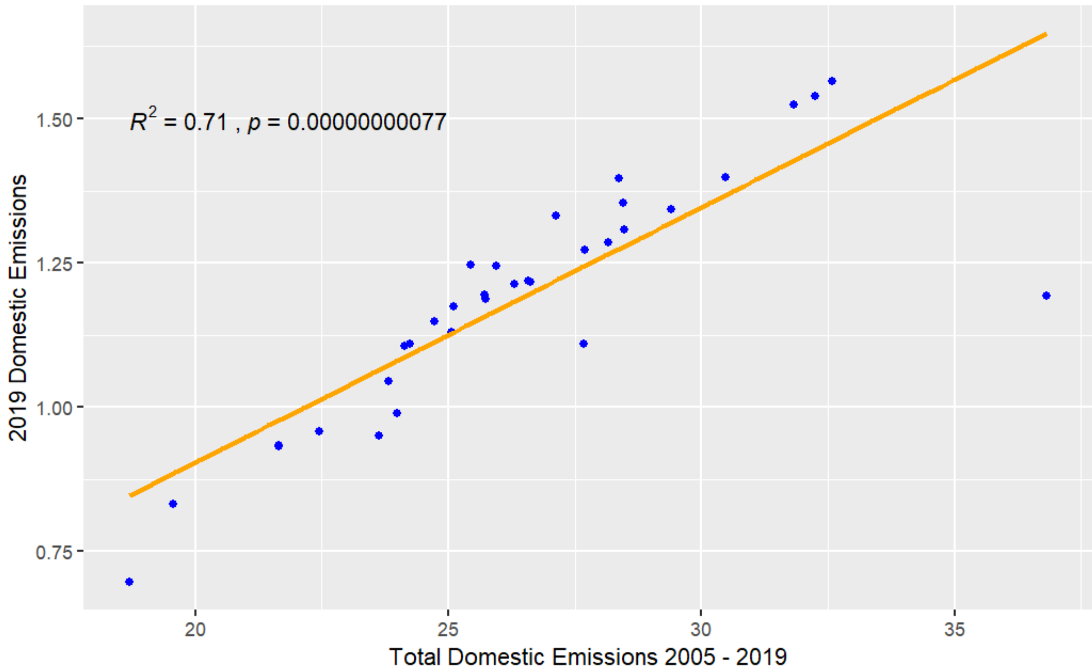
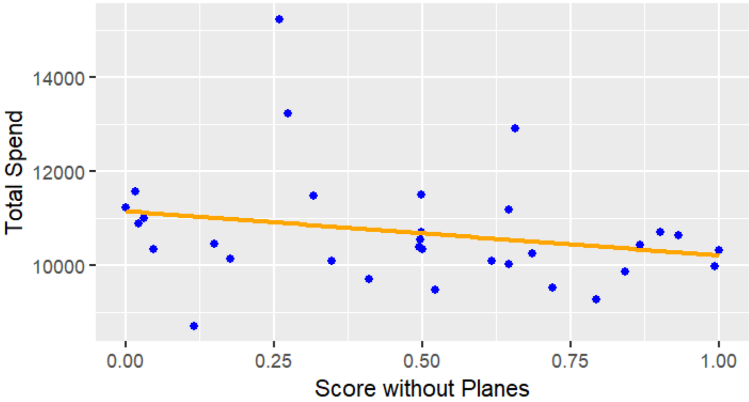


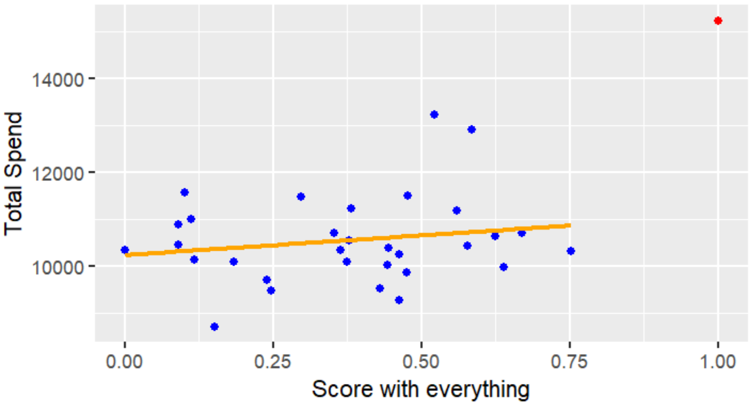
Figure 6.22: Domestic Emissions: 15 years vs 2019

LAD code	LAD	True CO2	Asthma Prevelence	Score	Score (with aeroplanes)	Score (without cars)
E09000001	City of London	NA	NA	NA	NA	NA
E09000002	Barking and Dagenham	22.44395	4.863193	0.79298	0.462752	0
E09000003	Barnet	30.48539	4.450571	0.647067	0.559254	0.708516
E09000004	Bexley	27.68778	5.261113	0.993728	0.63938	0.844953
E09000005	Brent	25.10688	4.354995	0.500809	0.364752	0.336144
E09000006	Bromley	32.25583	5.377503	0.900742	0.669565	0.707803
E09000007	Camden	23.63329	3.979714	0.016258	0.101595	0.372262
E09000008	Croydon	28.45765	5.045203	0.71996	0.430931	0.458953
E09000009	Ealing	25.4461	5.263195	0.499209	0.353523	0.319499
E09000010	Enfield	26.61573	4.899404	0.866944	0.577399	0.651441
E09000011	Greenwich	23.8198	4.858066	0.616728	0.375161	0.322133
E09000012	Hackney	21.64416	4.646648	0.048164	0	0.636549
E09000013	Hammersmith and Fulham	25.95051	4.04101	0.273457	0.521872	0.871994
E09000014	Haringey	25.7329	4.567348	0.348398	0.184732	0.516753
E09000015	Harrow	28.36721	5.341113	0.645821	0.44267	0.476702
E09000016	Havering	29.40396	5.229978	0.93222	0.624731	1
E09000017	Hillingdon	26.58207	5.255815	1	0.751882	0.654532
E09000018	Hounslow	24.73271	4.486989	0.494509	0.4451	0.352439
E09000019	Islington	23.99529	5.027855	0.030633	0.111879	0.448573
E09000020	Kensington and Chelsea	32.58789	3.378744	0.259752	1	0.779045
E09000021	Kingston upon Thames	28.16279	4.757316	0.497302	0.378658	0.798702
E09000022	Lambeth	24.23853	4.627001	0.177008	0.118107	0.246077
E09000023	Lewisham	25.06741	5.406817	0.411528	0.239613	0.483729
E09000024	Merton	27.13235	4.587985	0.499662	0.476529	0.472794
E09000025	Newham	19.56124	4.569632	0.115105	0.152359	0.554155
E09000026	Redbridge	25.71047	5.004048	0.685197	0.463683	0.854429
E09000027	Richmond upon Thames	31.82941	4.488499	0.657165	0.584585	0.607427
E09000028	Southwark	21.64253	4.53298	0.149788	0.091109	0.368031
E09000029	Sutton	28.48407	5.944153	0.842936	0.47576	0.542602
E09000030	Tower Hamlets	18.69193	4.533132	0.022894	0.090313	0.360306
E09000031	Waltham Forest	24.13289	5.060764	0.521426	0.247708	0.394314
E09000032	Wandsworth	26.29798	4.208996	0.317199	0.297572	0.753761
E09000033	Westminster	27.68277	3.358154	0	0.382512	0.343981

Table 6.5: Scores and Outcomes by LAD

Score	Coef	R-Squared	P-Value
Score without planes	-0.000059	0.06	0.14
Score with all	0.000037	0.03	0.34
score without planes or cars	0.000048	0.09	0.10





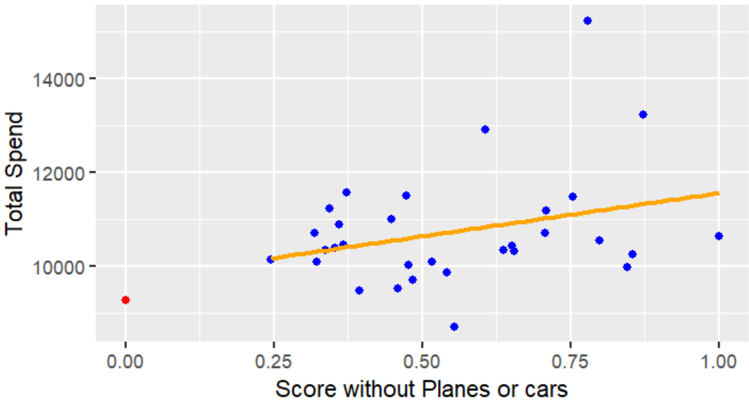


Figure 6.23: Scores vs Spend (LAD level)

Score	Coef	R-Squared	P-Value
Score	0.051	0.34	<0.001
Score (with aeorplanes)	0.061	0.64	<0.001
Score (without aeroplanes)	0.032	0.24	0.004

Table 6.6: Regression summary for Score ∼Asthma Prevalence models

Var	Coef	R-Squared	P-Value
Co2	0.0059	0.03	0.5
Asthma	-0.0046	0.65	<0.001

Table 6.7: Regression summary for Var $\sim$Flight Emissions

Figure 6.24: Average CO2 Emissions by LAD (not adjusted for baselines)

LAD	LAD name	Domestic Co2	Asthma Prevelence	Flight Emissions	Flyers Proporiton
E09000001	City of London	NA	NA	NA	NA
E09000002	Barking and Dagenham	22.44395	4.863193	99.68331	0.185246
E09000003	Barnet	30.48539	4.450571	232.4653	0.274764
E09000004	Bexley	27.68778	5.261113	111.0253	0.169989
E09000005	Brent	25.10688	4.354995	172.9975	0.241805
E09000006	Bromley	32.25583	5.377503	155.8553	0.196583
E09000007	Camden	23.63329	3.979714	270.859	0.307179
E09000008	Croydon	28.45765	5.045203	135.0167	0.196625
E09000009	Ealing	25.4461	5.263195	192.2183	0.255797
E09000010	Enfield	26.61573	4.899404	142.7009	0.220963
E09000011	Greenwich	23.8198	4.858066	146.5768	0.226253
E09000012	Hackney	21.64416	4.646648	212.0254	0.332664
E09000013	Hammersmith and Fulham	25.95051	4.04101	366.3182	0.370102
E09000014	Haringey	25.7329	4.567348	179.6689	0.275734
E09000015	Harrow	28.36721	5.341113	178.7535	0.213134
E09000016	Havering	29.40396	5.229978	141.5266	0.201803
E09000017	Hillingdon	26.58207	5.255815	177.2092	0.18807
E09000018	Hounslow	24.73271	4.486989	239.1083	0.259494
E09000019	Islington	23.99529	5.027855	270.4942	0.325173
E09000020	Kensington and Chelsea	32.58789	3.378744	605.4299	0.354658
E09000021	Kingston upon Thames	28.16279	4.757316	212.24	0.236803
E09000022	Lambeth	24.23853	4.627001	214.3741	0.305992
E09000023	Lewisham	25.06741	5.406817	167.5897	0.248494
E09000024	Merton	27.13235	4.587985	256.8261	0.259366
E09000025	Newham	19.56124	4.569632	255.8629	0.2588
E09000026	Redbridge	25.71047	5.004048	175.0262	0.236189
E09000027	Richmond upon Thames	31.82941	4.488499	245.4197	0.263364
E09000028	Southwark	21.64253	4.53298	207.6067	0.276981
E09000029	Sutton	28.48407	5.944153	120.0034	0.181337
E09000030	Tower Hamlets	18.69193	4.533132	264.5032	0.322596
E09000031	Waltham Forest	24.13289	5.060764	136.9945	0.241275
E09000032	Wandsworth	26.29798	4.208996	246.3445	0.314537
E09000033	Westminster	27.68277	3.358154	411.2499	0.284306

Table 6.8: Flights Vs Outcomes

	Gas Vehicle Owners	EV/Hybrid vehicle owners	No Vehicle
Flights	216335.5	435404.7	193821.7
Gas	575872.1	442448.2	8771.313
Car Ownership	201439	244095.9	0
TFL	29478.31	27690.04	50016.01
Clothes	297653.2	277080.7	279659.3
Other	8815.812	10997.67	13243.82

Table 6.9: Emissions from vehicle owners vs non-owners

LAD	LAD Name	Gas Emissions	TFL emissions	Taxi Emissions	E-bike Emissions
E09000001	City of London	NA	NA	NA	NA
E09000002	Barking and Dagenham	553.4817	159.3583	1.896561	0.080336
E09000003	Barnet	430.9938	169.8233	2.539035	0.136452
E09000004	Bexley	615.1416	116.8907	1.745565	0.068895
E09000005	Brent	376.7729	187.3128	2.914898	0.108924
E09000006	Bromley	570.3059	107.5567	1.608916	0.066915
E09000007	Camden	132.4476	204.9836	3.751843	0.170688
E09000008	Croydon	485.1837	146.4817	2.056476	0.084764
E09000009	Ealing	364.3346	158.8309	3.71954	0.14548
E09000010	Enfield	545.6748	141.8078	1.847353	0.079929
E09000011	Greenwich	433.9216	184.2069	2.114389	0.083262
E09000012	Hackney	136.7919	191.1657	3.531667	0.164549
E09000013	Hammersmith and Fulham	240.7687	197.4215	4.996664	0.197138
E09000014	Haringey	284.7617	214.7873	2.43103	0.123291
E09000015	Harrow	430.3192	148.8231	2.360256	0.103019
E09000016	Havering	579.365	132.819	2.326318	0.082862
E09000017	Hillingdon	625.9435	106.9171	2.084683	0.074823
E09000018	Hounslow	361.8495	146.783	2.835059	0.132016
E09000019	Islington	135.0386	185.5371	3.250071	0.179581
E09000020	Kensington and Chelsea	239.7013	138.0768	6.372591	0.180414
E09000021	Kingston upon Thames	342.5433	120.5468	2.192047	0.102712
E09000022	Lambeth	205.6417	218.5921	3.396444	0.148932
E09000023	Lewisham	313.8678	196.0841	2.007352	0.099298
E09000024	Merton	358.1595	184.1574	2.286264	0.108945
E09000025	Newham	176.6351	254.9021	2.826981	0.106166
E09000026	Redbridge	447.9235	171.6703	2.283858	0.089052
E09000027	Richmond upon Thames	429.7522	113.701	2.346171	0.126751
E09000028	Southwark	190.0049	199.6729	2.688604	0.14211
E09000029	Sutton	537.5241	99.6665	1.835942	0.075763
E09000030	Tower Hamlets	139.3744	221.7795	3.295356	0.130937
E09000031	Waltham Forest	378.2747	195.007	2.013545	0.096483
E09000032	Wandsworth	258.0531	202.066	3.407004	0.155875
E09000033	Westminster	113.4874	405.0382	3.77967	0.133247

Table 6.10: Average emissions per person from Intra-London Travel

Model	Coef	R-Squared	P-Value
Co2 ~ Gas	0.016	0.24	0.004
Asthma ~ Gas	0.0034	0.4	<0.001
Co2 ~ Ev	98	0.13	0.045
Asthma ~ Ev	-25	0.3	0.001

Table 6.11: Fuel emissions vs Outcomes Variables coefficients

LAD	LAD Name	Clothes Emissions (kg)	Food Rescue Emissions (kg)	Organic Produce Transactions	Local Spend Transactions
E09000001	City of London	NA	NA	NA	NA
E09000002	Barking and Dagenham	193.4711	-0.19672	0	0.070492
E09000003	Barnet	316.9867	-0.38396	0.07704	0.157547
E09000004	Bexley	343.4737	-0.43349	0.033458	0.108373
E09000005	Brent	252.2937	-0.39905	0.059895	0.235629
E09000006	Bromley	319.5868	-0.30503	0.078151	0.112156
E09000007	Camden	258.2776	-0.5793	0.118331	0.343489
E09000008	Croydon	274.3719	-0.2876	0.068259	0.140377
E09000009	Ealing	244.4001	-0.31252	0.112904	0.116661
E09000010	Enfield	308.8583	-0.38275	0.05824	0.201763
E09000011	Greenwich	253.6058	-0.44855	0.072342	0.156332
E09000012	Hackney	303.9364	-0.49337	0.148895	0.239052
E09000013	Hammersmith and Fulham	339.1847	-0.41993	0.085931	0.134494
E09000014	Haringey	286.558	-0.46709	0.171063	0.231185
E09000015	Harrow	275.8169	-0.25313	0.052233	0.198806
E09000016	Havering	366.6592	-0.28527	0.012449	0.141458
E09000017	Hillingdon	308.6329	-0.37614	0.033172	0.175439
E09000018	Hounslow	254.3087	-0.36027	0.067201	0.075949
E09000019	Islington	270.9749	-0.40963	0.118919	0.186047
E09000020	Kensington and Chelsea	316.6963	-0.31553	0.11577	0.262733
E09000021	Kingston upon Thames	335.9858	-0.47755	0.106438	0.22003
E09000022	Lambeth	236.694	-0.40605	0.143409	0.105294
E09000023	Lewisham	283.6789	-0.56627	0.168726	0.212952
E09000024	Merton	280.6853	-0.5105	0.126171	0.102511
E09000025	Newham	290.9695	-0.40258	0.044189	0.275657
E09000026	Redbridge	341.7083	-0.23859	0.048779	0.136509
E09000027	Richmond upon Thames	299.9168	-0.31638	0.215498	0.240765
E09000028	Southwark	264.3056	-0.62978	0.187777	0.118852
E09000029	Sutton	290.9543	-0.40424	0.053321	0.159739
E09000030	Tower Hamlets	255.796	-0.37692	0.064462	0.128846
E09000031	Waltham Forest	267.001	-0.53081	0.091663	0.084446
E09000032	Wandsworth	324.9705	-0.54941	0.105631	0.166839
E09000033	Westminster	250.6732	-0.35004	0.042613	0.105177

Table 6.12: Consumer Purchases

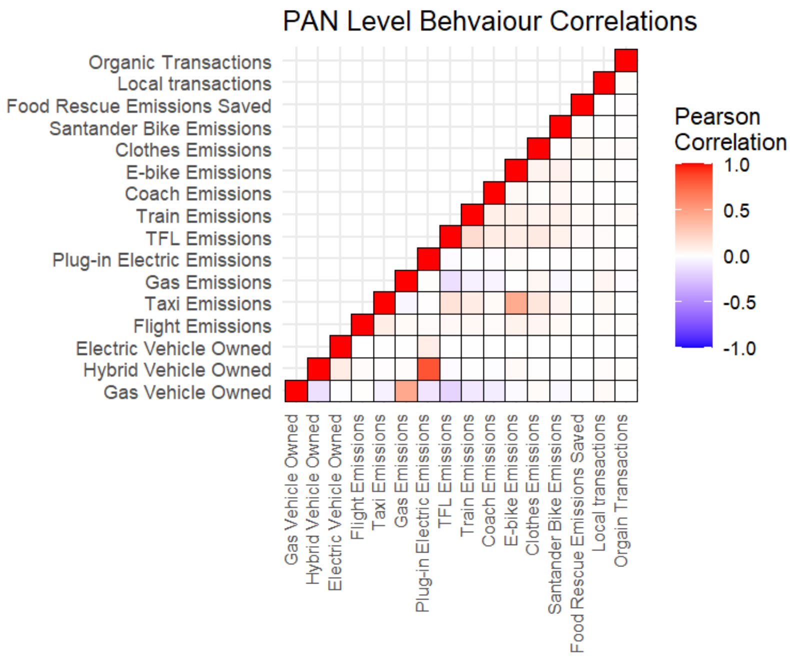


Figure 6.25: Correlation between behaviours (PAN level)

Github access available upon request.

Bibliography

[1] Gary Kerr. Climate change: The role of the individual. *Scottish Parliament Information Centre Research Briefing*, 07 2012.

[2] Visa Inc. Overview of payments card product solutions. https://partner.visa.com/site/learn/overview-of-payments.html. (accessed: 12.08.2022).

[3] UN. The paris agreement. https://www.un.org/en/climatechange/paris-agreement: :text=Today%2C%20193%20Parties%20(192%20countries,strengthen%20their%20commitments%20ov (accessed: 21.07.2022).

[4] The World Bank. Poverty and climate change. 2009.

[5] Georgina Kyriacou Bob Ward. How is climate change affecting river and surface water flooding in the uk? 2021.

[6] Press Office. Record breaking temperatures for the uk. https://www.metoffice.gov.uk/about-us/press-office/news/weather-and-climate/2022/red-extreme-heat-warning-ud. (accessed: 21.07.2022).

[7] William James Kate Holton. Britain counts cost of historic heatwave as 13 die. https://www.reuters.com/world/uk/britain-counts-cost-hottest-day-ever-2022-07-20/. (accessed: 10.08.2022).

[8] Ipcc, 2018: Global warming of 1.5°c. an ipcc special report on the impacts of global warming of 1.5°c above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [v. masson-delmotte, p. zhai, h. o. pörtner, d. roberts, j. skea, p.r. shukla, a. pirani, w. moufouma-okia, c. péan, r. pidcock, s. connors, j. b. r. matthews, y. chen, x. zhou, m. i. gomis, e. lonnoy, t. maycock, m. tignor, t. waterfield (eds.)]. in press.

[9] UN Environemnt Programme. Fashion's tiny hidden secret. https://www.unep.org/news-and-stories/story/fashions-tiny-hidden-secret. (accessed: 14.08.2022).

[10] Laura Paddison. How carbon might go out of fashion. https://www.bbc.com/future/article/20211105-how-carbon-might-go-out-of-fashion. (accessed: 14.08.2022).

[11] Food and Drink Federation. Achieving net zero: A handbook for the food and drink sector. 2021.

[12] Mastercard. Turn purchases into meaningful action. https://www.mastercard.us/en-us/vision/corp-responsibility/priceless-planet/carbon-calculator.html. (accessed: 27.07.2022).

[13] Statista. Number of credit cards and debit cards per capita in the united kingdom from 2010 to 2020. https://www.statista.com/statistics/1098929/credit-cards-and-debit-cards-per-capita-in-the-united-kingdom/: :text=Credit%20cards%20and%20debit%20cards,the%20United%20Kingdom%202010%2D202 (accessed: 27.07.2022).

[14] Visa Europe. Principles for financial market infrastructures. 2019.

[15] Visa Inc. Visa fact sheet. https://usa.visa.com/dam/VCOM/global/about-visa/documents/aboutvisafactsheet.pdf. (accessed: 12.08.2022).

[16] Visa Inc. Visanet catalyst for commerce. https://usa.visa.com/content/dam/VCOM/download/corporate/media/technology/VisaNet-Network-Processing-Overview.pdf. (accessed: 12.08.2022).

[17] Adam Eric Greenberg and Hal E. Hershfield. Financial decision making. *Consumer Psychology Review*, 2(1):17–29, 2019.

[18] Shigeru Matsumoto. The opportunity cost of pro-environmental activities: Spending time to promote the environment. *Journal of Family and Economic Issues*, 35, 03 2013.

[19] Office for National Statistics. Estimates of the population for the uk, england and wales, scotland and northern ireland. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets (accessed: 12.08.2022).

[20] Rebecca Cafe. Oxford street: Wooing shoppers back to europe's busiest high street. 2021.

[21] City of London Corporation. Our role in london. https://www.cityoflondon.gov.uk/about-us/about-the-city-of-london-corporation/our-role-in-london: :text=In%20just%201.12%20square%20miles,Chancery%20Lane%20and%20Liverpool%20Street. (accessed: 12.08.2022).

[22] Greater London Authority. Statistical gis boundary files for london. https://data.london.gov.uk/dataset/statistical-gis-boundary-files-london. (accessed: 12.08.2022).

[23] Gov.uk. Uk greenhouse gas emissions statistics. https://www.gov.uk/government/collections/uk-greenhouse-gas-emissions-statistics. (accessed: 12.08.2022).

[24] Office for Health Improvment and Disparity. Asthma: Qof prevalence (all ages) - retired after 2019/20 (now 6+ yrs). https://fingertips.phe.org.uk/search/asthmapage/4/gid/1/pat/159/par/K02000001/ati/15/are/E92000001/iid/2 1/ctp/-1/yrr/1/cid/4/tbm/1. (accessed: 12.08.2022).

[25] Energy Industrial Strategy Department for Business. Uk local authority and regional carbon dioxide emissions national statistics: 2005 to 2019. https://www.gov.uk/government/statistics/uk-local-authority-and-regional-carbon-dioxide-emissions-national-statistics-2005-to-2019. (accessed: 12.08.2022).

[26] Dr. N.E. Ligterink. Dutch market fuel composition for ghg emissions. 2016.

[27] Energy Industrial Strategy Department for Business. Energy prices road fuels and other petroleum products. https://www.gov.uk/government/statistics/weekly-road-fuel-prices. (accessed: 10.08.2022).

[28] Petrol Retailers Association. The uk retail road fuels market 2020. https://www.ukpra.co.uk/en/about/facts-and-figures. (accessed: 1.08.2022).

[29] Department for Transport. Average new car and light goods vehicle (lgv) fuel consumption: Great britain, 1997-2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1040512/en (accessed: 8.08.2022).

[30] Asda. Pay at pump is changing. https://www.asda.com/payatpump. (accessed: 22.08.2022).

[31] Calum Edmunds Keith Bell James Dixon, Waqquas Bukhsh. Scheduling electric vehicle charging to minimise carbon emissions and wind curtailment. 2020.

[32] RAC. Cost to rapid charge an electric car rises by a fifth in eight months. https://www.rac.co.uk/drive/news/electric-vehicles-news/cost-to-rapid-charge-an-electric-car-rises-by-a-fifth-in-eight-months1/. (accessed: 1.08.2022).

[33] Rona Webster. Can the electricity distribution network cope with an influx of electric vehicles? *Journal of Power Sources*, 80(1):217–225, 1999.

[34] Low Carbon Vehicle Partnership. Lifecycle emissions from cars. 2015.

[35] SMMT. Average vehicle age. https://www.smmt.co.uk/industry-topics/sustainability/average-vehicle-age/, 2022.

[36] Transport for London. How many cars are there in london and who owns them? *IOP Conference Series: Earth and Environmental Science*, 2019.

[37] TfL. Pay as you go caps. https://tfl.gov.uk/fares/find-fares/tube-and-rail-fares/pay-as-you-go-caps. (accessed: 28.08.2022).

[38] TFL. Safety, health and environment annual report. https://content.tfl.gov.uk/safety-health-and-environment-report-2020-21.pdf. (accessed: 1.08.2022).

[39] TFL. How we are funded. https://tfl.gov.uk/corporate/about-tfl/how-we-work/how-we-are-funded. (accessed: 1.08.2022).

[40] TFL. Carbon footprint of a tube journey. https://tfl.gov.uk/corporate/transparency/freedom-of-information/foi-request-detail?referenceId=FOI-0354-2021. (accessed: 8.08.2022).

[41] Department for Transport. Taxi and private hire vehicle statistics, england: 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/997793/tax and-private-hire-vehicle-statistics-2021.pdf. (accessed: 30.07.2022).

[42] Patrick Jochem Patrick Plötz, Simon Árpád Funke. Empirical fuel consumption and co2 emissions of plug-in hybrid electric vehicles. 2017.

[43] Thomas Kingsley. Uber increases london fares for first time since 2017. https://www.independent.co.uk/news/uk/uber-london-fare-hike-drivers-b1955963.html. (accessed: 10.08.2022).

[44] TFL. Taxi fares. https://tfl.gov.uk/modes/taxis-and-minicabs/taxi-fares. (accessed: 10.08.2022).

[45] Kay W.Axhausen Daniel J.Reck, Henry Martin. Mode choice, substitution patterns and environmental impacts of shared and personal micro-mobility. 2022.

[46] BikeRadar. How green is cycling? riding, walking, ebikes and driving ranked. https://www.bikeradar.com/features/long-reads/cycling-environmental-impact/. (accessed: 2.08.2022).

[47] Transport for London. Foi request. https://tfl.gov.uk/corporate/transparency/freedom-of-information/foi-request-detail?referenceId=FOI-4535-1718. (accessed: 14.08.2022).

[48] TFL. Santander cycles membership. https://tfl.gov.uk/modes/cycling/santander-cycles/register-for-your-own-key?intcmp=2312. (accessed: 2.08.2022).

[49] Energy Industrial Strategy Department for Business. 2020 government greenhouse gas conversion factors for company reporting. 2020.

[50] Department for Transport. Rail factsheet. 2019.

[51] Morten Birkvedc Tina Joanesd Wencke Gwozdzde Joshua Sohna, Kristian S.Nielsenb. The environmental impacts of clothing: Evidence from united states and three european countries. 2021.

[52] Shipping Storm. Approximate weight of goods. https://shippingstorm.com/en/list-of-weight/. (accessed: 19.08.2022).

[53] Too Good To Go. The carbon footprint. https://toogoodtogo.com/en-us/movement/knowledge/the-carbon-footprint. (accessed: 2.08.2022).

[54] OddBox. Meet our fruit and veg boxes. https://www.oddbox.co.uk/boxes/fruit-and-veg-box. (accessed: 2.08.2022).

[55] Lisa Elaine Held. The real climate impact of organic farming. 2029.

[56] Hertel T. Sampson G. Avetisyan, M. Is local food more environmentally friendly? the ghg emissions impacts of consuming imported versus domestically produced food. *Environmental and Resource Economics*, 58:415–462, 2014.

[57] Paul e. pfeffer, ian s. mudway, jonathan grigg, air pollution and asthma: Mechanisms of harm and considerations for clinical interventions, chest, volume 159, issue 4, 2021, pages 1346-1355, issn 0012-3692, https://doi.org/10.1016/j.chest.2020.10.053.

[58] Gordon Macleod. The grenfell tower atrocity: Exposing urban worlds of inequality, injustice, and an impaired democracy. *City*, 22:460–489, 07 2018.

[59] Hayley Bennett. Distribution of individual total wealth by characteristic in great britain: April 2018 to march 2020. 2022.

[60] Transport for London. Ulez and lez services. https://tfl.gov.uk/modes/driving/lez-lez-services-37309. (accessed: 22.08.2022).

[61] Joshua S. Graff Zivin, Matthew J. Kotchen, and Erin T. Mansur. Spatial and temporal heterogeneity of marginal emissions: Implications for electric cars and other electricity-shifting policies. *Journal of Economic Behavior Organization*, 107:248–268, 2014.

[62] EE News Sara Schonhardt. Rich americans have higher carbon footprints than other wealthy people. 2021.

[63] EMILY GHOSH SIVAN KARTHA, ERIC KEMP-BENEDICT, STOCKHOLM ENVIRONMENT INSTITUTE ANISHA NAZARETH, and OXFAM TIM GORE. The carbon inequality era. 2020.

[64] Business Department for Energy and Industrial Strategy. 2020 uk greenhouse gas emissions, final figures. https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions: :text=Transportation%20(27%25%20of%202020%20greenhouse,ships%2C%20trains%2C%20and%20pla (accessed: 14.08.2022).

[65] European Central Bank. Study on the payment attitudes of consumers in the euro area (space). 2020.