Estimating geographical retail markets from card spending data

Samir Doshi* Vicky Hoolohan† Tabitha Lewis ‡ Jakob Schneebacher§

September 30, 2024

Abstract

Accurate market definitions are important for competition agencies, but traditional survey-based measures are costly, time-consuming and noisy at low aggregations. This paper explores the use of consumer card spending data to improve the timeliness and accuracy of retail market estimates. With the help of a standard machine-learning algorithm, we cluster spending flows from cardholder postcode sectors to merchant postcode sectors for detailed categories of retail merchants in the UK at a monthly frequency. To decide the thresholds for the clustering algorithm, we use estimates of average distance travelled from traditional survey tools. We find geographical retail markets that differ systematically by merchant good category and across space. Market size is also predicted by demographic and economic characteristics. Over time, market size is relatively stable but shrinks during periods of pandemic-induced travel restrictions. Markets for different retail goods are spatially correlated in predictable ways. Beyond applications to competition agency casework, this method allows researchers to investigate local competition and the impact of technology and government policies on spatial consumer search and purchasing behaviour.

JEL Classification: D40; L10; L81; R12
Keywords: market definitions; consumer behaviour; real-time data; economic data science

*Competition and Markets Authority (CMA).

†Office for National Statistics (ONS).

‡Office for National Statistics (ONS) and Swansea University.

§Competition and Markets Authority (CMA), King's College London (KCL) and Economics Statistics Centre of Excellence (ESCoE). We are grateful to Jonatan Benarroch, Hannah Brown, Tom Farmer, Joel Kariel, Keith Lai, Simon Martin, Howard Smith, Joel Stiebale, Dayal Strub and Mike Walker for helpful feedback. Any views expressed are solely those of the authors and cannot be taken to represent those of the CMA, ONS or Visa. Corresponding author: Jakob Schneebacher (jakob.schneebacher@cma.gov.uk).

1 Introduction

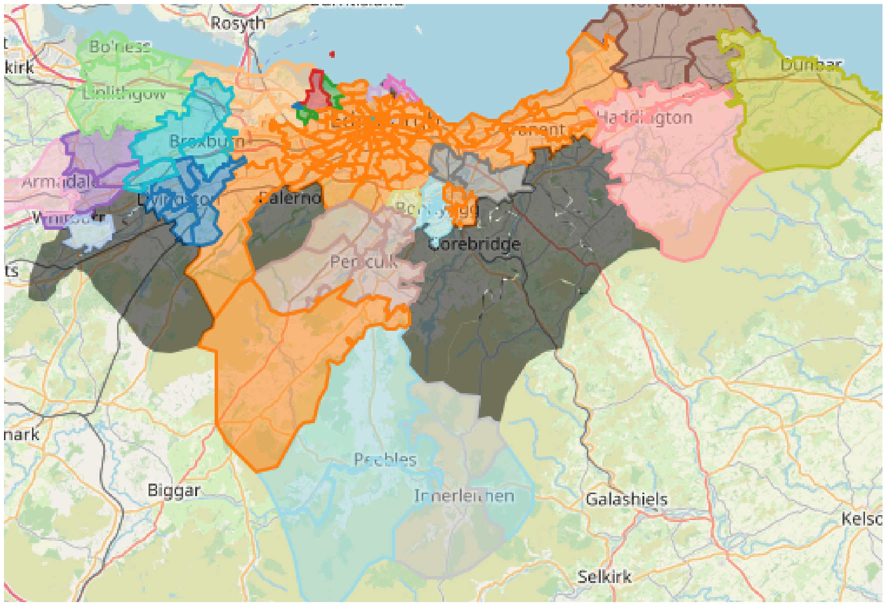
Competition agencies frequently make use of market definitions to understand firm conduct and market power, and to estimate the likely impact of proposed mergers. But measuring markets accurately is difficult. For the aggregate economy, economists often define markets from Standard Industrial Classification (SIC) codes, but this has two drawbacks. First, SIC industries are defined from production technologies and are therefore simultaneously too narrow and too broad compared to the choice sets of consumers. Second, they do not capture the important geographical dimension of many markets. In competition enforcement work (like market studies and merger cases) on the other hand, agencies often use customer surveys for market definitions. But surveys are costly, time-consuming and noisy. They also only capture a snapshot of a market at one point in time.

We offer an alternative approach, which makes use of consumer card spending data in narrowly defined merchant categories. We apply this method to estimate geographical retail markets in the UK. To do so, we use card payment flows from the location associated with consumers who make card payments to the location associated with the merchant who receives it, within a so-called Merchant Category Group (MCG). For disclosure control reasons, locations are reported at small geographies called postcode sectors. MCGs define groups of similar merchants, such as traditional sit-down restaurants, quick-service restaurants or apparel stores.

We apply a simple clustering algorithm to group postcode sectors together such that consumers within each cluster shop at the same merchants for a given MCG, and merchants within a cluster serve the same consumers. We call these clusters "local retail markets" for a given MCG. This approach captures the intuition that firms within a market are competing for the same customers, and customers within a market choose between the same firms. It also results in more realistic, location-based markets than for instance a SIC-based market definition that implicitly defines all markets as national, and more complete and time-varying market definitions when compared to traditional survey methods. Given sufficient data, this method in principle also allows researchers to define separate markets for different quantiles of the card spending distribution, therefore taking into account vertical price differentiation of goods and services within MCGs.

To fix ideas, Figure 1 shows the estimate of our local markets for traditional sit-down restaurants for the city of Edinburgh in March 2024. Geographical contiguity of these markets is not imposed by the estimation procedure, but rather reflects consumer choices, where it emerges. Likewise, the geographical size of the markets reflects transport links and consumer travel behaviour. Figure

Figure 1: Cluster locations across Edinburgh, restaurants, March 2024



2 plots market-level merchant counts, across postal sectors, in traditional sit-down restaurants in March 2024 in Great Britain. The map shows distinct differences in market size and merchant counts between rural and urban markets.

The estimated market definitions have two applications. First, for competition agencies like the UK Competition and Markets Authority (CMA), the estimates represent a helpful ready-to-use tool for daily casework. For instance, in many merger cases the size of the relevant local market is an important question. This new tool can be used to complement or cross-validate estimates obtained via traditional survey-based methods. Estimating local market size can be expensive in terms of time and resources when case teams have to rely on customer data to compute the catchment areas.[1]

The paper makes a methodological and a substantive contribution. Methodologically, we provide a new approach to obtain estimates of local market size and shape by MCGs across the UK. This methodology allows markets to vary flexibly in size and shape across locations, by clustering areas based on card spending flows from cardholders to merchants. This represents an improvement over current local market estimates in casework, where the area under investigation is often divided into equally-sized local "markets". While our estimates cannot replace case-specific efforts to define local markets (which make use of product-specific and often proprietary data), they can be used

[1]That is, the "area from which a provider is expected to draw the vast majority of its business". See for instance https://www.oxera.com/insights/agenda/articles/catchment-if-you-can-a-practical-guide-to-local-competition-analysis/.

as case background, and to ground casework estimates. Since the market definitions in this paper are based on existing card spending data, market definitions can be computed retroactively, and at high frequencies.

Substantively, our market estimates enable a better understanding of patterns in consumer search and consumption behaviour, as captured in travel distances and spending amounts for retail shopping purposes across the UK, and of local competition in retail markets. We find geographical retail markets that differ systematically by MCG and across space. Market size is also predicted by demographic and economic characteristics. Over time, market size is relatively stable but shrinks during periods of pandemic-induced travel restrictions. Markets for different retail goods are spatially correlated in predictable ways. For instance, markets for food and grocery retail are least correlated with other MCGs, perhaps reflecting that most UK consumers buy groceries locally. Apparel and accessories market are highly correlated with discount stores, home improvement stores and entertainment venues, perhaps capturing suburban shopping malls. Geographically, markets are smaller in cities and larger for postcodes with larger car ownership shares. There are important regional differences too. These findings suggest important improvements are possible over nationally-uniform, distance-based local market estimates currently used in competition casework.[2]

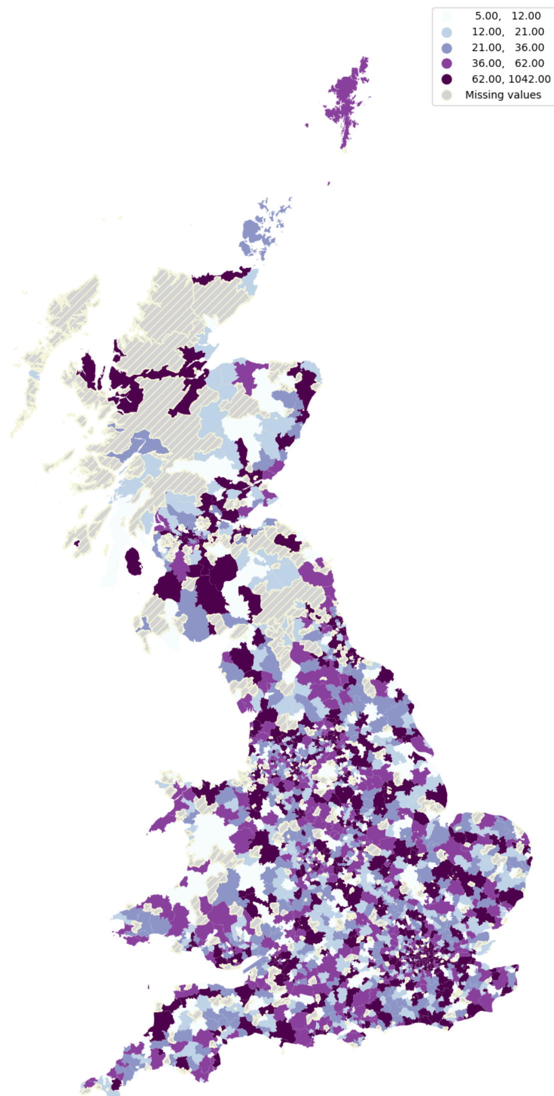
In this article, we highlight three immediate policy-relevant applications. First, our method allows us to understand how geographic retail markets in the UK have changed over time (including during the Covid-19 pandemic). Second, the degree of spatial correlation between consumers' retail spending patterns across MCGs gives insight into joint purchase decisions and their drivers, including the product portfolio choices of retailers. Finally, by providing merchant counts within each local market, we provide a new measure of local competition in UK retail markets.

Alongside this paper, we are making our local market estimates available to the research community. We hope this enables researchers to link other data sources to our market estimates, and answer many further research questions. For instance, by combining this data with local price data, we may get a better understanding of market structure and firm conduct in non-tradable goods and services, similar to existing studies of the impact of local competition on consumer outcomes in specific markets with well-understood boundaries, for instance road fuel (Martin, 2020; CMA, 2023a; Byrne et al., 2024).

Estimates of market concentration have played a large role in the recent debate on the causes and

[2]By defining markets for quantiles of the card spending distribution, research can additionally use this methodology to obtain a better picture of inequalities in market access across the income distribution, including for instance to measure the existence of food deserts CMA (2023b).

Figure 2: Cluster locations across Great Britain (including counts of merchants per cluster), restaurants, March 2024



consequences of aggregate market power (Crouzet and Eberly, 2019; Kwon et al., 2024; White and Yang, 2020), triggered by the seminal paper by De Loecker et al. (2020). Covarrubias et al. (2020) for instance argue that while US concentration in the aggregate had risen for benign, productivity-enhancing reasons in the past, more recent changes in concentration are driven by anti-competitive conduct. However, as Shapiro and Yurukoglu (2024) argue, concentration measures can be misleading if markets are defined incorrectly. In addition to defining markets over products or business lines rather than industries, this implies the need to get the geographical scope of markets right. Recent papers Autor et al. (2023) and Rossi-Hansberg et al. (2021) find diverging national and local concentration trends for employment and turnover respectively, when defining markets as small US Census geographies. Hsieh and Rossi-Hansberg (2023) provide a theory to rationalise these diverging trends, linked to rising returns to scale in services. Benkard et al. (2021) for the US and Calligaris et al. (2024) for Europe likewise show that aggregate concentration trends look dramatically differently when markets are defined over products and at different geographical scales. Perhaps closest to this paper in this literature is Patterson and Vavra (2024), who use US card spending data on both in-person and online purchases to characterise market concentration at a consumer location. Using market boundaries based on actual consumer behaviour represents a further step towards understanding secular trends in market power, and when combined with administrative data on the location of local establishments (Lui et al., 2023) these estimates may enable a better understanding of the firm location decisions and competitive dynamics behind the aggregate trends.

Within industrial economics, economists have over time proposed various tools to define geographic markets for specific goods. Audy and Erutku (2005) apply price correlation and Granger causality tests across regions to define geographic markets in the Canadian gasoline industry. Elzinga and Hogarty (1973) propose the use of product shipment patterns to define geographic markets. However, Elzinga and Swisher (2011) identify some limitation of this approach. They point out problems when consumers, rather than goods and services, move around. Additionally, the data requirements to apply this test are difficult to meet for a variety of industries. Ulrick et al. (2020) also propose an approach for defining geographic markets under less stringent information requirements, using consumer search behaviour. Eizenberg et al. (2021) use local price variation across neighbourhoods in the city of Jerusalem to understand local markets and price segmentation. Gibbs et al. (2024) provide a guide to using consumer credit report data to answer a host of economic questions, including consumer mobility and purchases of goods often made on credit, such as cars.

Methodologically, the paper closest in spirit to the present one is Batch et al. (2023). Batch et al.

(2023) use US county-level spending data to group counties into "Consumption Zones". By applying the standard methodology used to create "Commuting Zones" (Tolbert and Sizer, 1996) in a labour market context, and using county sales flow data from Fiserv (a card transaction intermediary), the authors create Consumption Zones, which they take to represent local consumption markets. Batch et al. (2023) use a hierarchical clustering approach that applies a dissimilarity matrix to identify how similar one county is to another, based on the proportion of total spending flows between counties. The hierarchical clustering algorithm then groups counties together based on that score. We improve on this paper in a few ways. First, by using much more granular spatial data, and detailed retail merchant categories, we come much closer to local market estimates obtained via traditional methods. Second, we do not impose the same symmetry assumption, allowing merchants in a local market to be located in fundamentally different postcode sectors than consumers, which we believe more realistically reflects spatial sorting. Finally, we use market-specific survey estimates of distance travelled to choose thresholds that realistically reflect consumer behaviour. To distinguish the symmetric county-level aggregate consumption flow clusters from our more granular, retail merchant type specific and potentially asymmetric clusters, we call the latter "local retail markets".

The rest of the paper is organised as follows. Section 2 describes the card payments data we use. Section 3 explains the clustering algorithm, with particular attention to the threshold selection. Section 4 shows the characteristics of the resulting market estimates, maps them out across space, and provides some initial analysis. A brief final section discusses further applications and concludes.

2 Data

2.1 Consumer card spending data

To construct local retail market estimates, we use payment card spending data provided by Visa, a large payment network. This data only captures one of the possible payment channels used by UK consumers. Visa aggregates and anonymises customer card transaction data before sharing the dataset with the UK Office for National Statistics (ONS). Spending flows are aggregated to the merchant postal (also referred to as postcode) sector level and the cardholder postal sector level. Postcode sectors are derived from postcodes. UK postcodes generally take the form XXNN NXX (where "X" stands in for letters and "N" for digits). Postal sectors consist of the first one or two letters of the postcode (the postcode area), a one or two digit code (which, combined with the postcode area, gives the postcode district) and finally the first character of the inward location code (the part of UK postcodes after the space). For example, the UK postcode DL3 7EE would be associated with the postal sector DL3 7.[3]

Using postal-sector level data is a key contribution of this work, and brings the general methodology of Batch et al. (2023) much closer to actual local market definitions. There are over 12,000 postal sectors in the UK alone (ONS, 2023c), compared to 3,128 counties across the whole US in Batch et al. (2023). UK markets tend to be much smaller than UK counties. For instance, the CMA's merger analysis in the groceries sector, in the case of Sainsury's/Asda, primarily considered a 15-minute drive-time catchment area CMA (2017). The median area of a postal sector for comparison is 2 square miles. This allows us to define sufficiently granular retail markets through spatial consumer spending patterns.[4]

We observe card spending monthly and at the merchant category group (MCG) level, from January 2019 to June 2024. MCGs, of which there are currently 25 in total, capture retail merchants supplying similar goods or services. For instance, traditional sit-down restaurants and quick-service restaurants (QSR) are separated into different MCGs as they likely do not compete directly with each other. On the other hand, sit-down restaurants are grouped together regardless of the cuisine they serve. More information about MCGs can be found in appendix A.

To give a sense of the data coverage, taking the month of June 2023 and selecting the food and grocery MCG, the dataset contained spending flows from consumers in 8,337 postal sectors to 3,938 merchant postal sectors that month. The difference between these numbers and the previously-mentioned 12,000 postal sectors is accounted for by (1) postcodes that do not feature cardholders or merchants, (2) postcodes that do not show transactions in this particular month and (3) postcodes with small enough transaction counts or spending flows to be dropped during statistical disclosure control in order to preserve the anonymity of individual cardholders.

The dataset contains the number of transactions, total spend, and total number of cardholders from one postal sector to another, at the MCG-by-month level, exclusively for face-to-face transactions. We then combine our baseline dataset on card spending between postal sectors by MCG with geospatial data from the Office for National Statistics (ONS). This information is based on ONS (2016). We calculate distances between postal sector centroids and therefore assign a distance measure of 0 where spending takes place within the same postal sector. We further augment this data with postal sector characteristics from the decennial UK Census (ONS, 2024a).

[3] The corresponding postal district would be DL3, and the postal area would be DL.

[4] To find out more detail about the data and the initial ONS analysis, see ONS (2023b) for a summary of the data, and ONS (2023a) for descriptive statistics.

Visa card spending data has three key benefits for this research project. First, the quality and coverage of the data is higher than for other, comparable data sources. Additionally, card spending makes up the majority of UK transactions. By contrast, the UK Payment Systems Regulator estimates that cash accounted for only 23% of transactions in the UK, down from 58% ten years earlier.[5] Second, the data records transactions at a monthly level and is updated every quarter, enabling close to real-time analysis. The data series also begins in 2019, allowing us to analyse local retail markets before, during, and after the Covid-19 pandemic, and the more recent UK cost-of-living crisis. Third, the high spatial resolution at the postal sector level, subject to appropriate dominance and statistical disclosure suppression, provides a unique dataset with high geographical resolution. This enables us to define appropriately small local markets.

However, the data also has limitations. First, many observations are excluded in our data to comply with Visa's disclosure control procedures. An observation may be excluded based on Visa's data confidentiality and data privacy standards if a small number of merchants accounts for a large share of spending. While coverage varies across MCGs, our dataset nevertheless contains a substantial share by value of all Visa transactions. Additionally, Visa also apply ad-hoc exclusions for further disclosure purposes.

Second, cardholder locations are inferred by Visa, based on the transaction flows of the cardholder. The cardholders' home address is not known by Visa. This means that, if the cardholder is temporarily based in another location (for instance, on holiday) in a given month or makes most of their transactions around their workplace, the cardholder location may not correspond to their home address. However, given that we are interested in areas from which local merchants draw their customer base, we believe this is the right definition of cardholder location for the purpose of defining local markets.

Third, the classification of merchants is not straightforward. A merchant is classified based on their primary business, but they could sell many things at once. For example, a large supermarket would be classed as food and grocery but could also sell clothes and electronics.[6] Additionally, if a department store with different counters sells distinct categories of consumer goods, then each counter may be classified as a separate merchant. In some cases, each merchant may correspond to a distinct brand, although this is unlikely to be the case in our postal sector level. In most cases, each merchant would simply correspond to a distinct retail shop.[7] For our purposes, MCGs are

[5]Source: https://www.psr.org.uk/media/20ob5wee/payments-over-time.pdf
[6]Note however that department stores are a separate MCG.
[7]Online retailers are excluded altogether from this dataset.

therefore a more precise classification than Standard Industrial Classification (SIC) codes, but as more data becomes available, finer MCG groupings will become possible. This would in principle allow researchers to construct their own merchant groupings by aggregating MCGs believed to be substitutable amongst each other for the purposes of the research question at hand.

Fourth, macroeconomic shocks and seasonality may matter when considering the analysis across time. There were far fewer transactions during the Covid-19 pandemic, which may have led to anomalous exclusions, and there is a significant seasonal component in the data to reflect consumers' spending patterns, with a clear "holiday pattern" emerging in some MCGs during the summer.

Finally, despite the excellent coverage and the widespread use of payment cards in the UK, the population of cardholders may differ systematically from the wider population of consumers. Demographic and economic factors may influence the usage of Visa cards in comparison to other payment methods, which could have some impact on the coverage of transactions. For example, income and consumer behaviour may influence payment mode (such as use of credit or debit card). Additionally, usage of cash in comparison to cards may be different across different demographics. In 2019, adults over the age of 55 years were twice as likely to pay by cash than those under 35 PSR (2021). Some merchants may not accept cards, and these may likely be disproportionately small businesses.

2.2 Merchant data

For certain parts of the analysis, we compute measures of merchant counts and merchant concentration within local retail markets. Visa provide aggregated data containing only the total count of merchants, separately from our main dataset described above. This additional dataset contains the spend, number of transactions, and number of distinct merchants per month and quarter. It is aggregated by merchant postal sector. Data suppression for disclosure purposes is still applied, as described in Section 2.1 .

We link this dataset on the postcode sector level to our primary cardholder spending dataset to infer the number of merchants in each market. However, there are minor differences with how spending is attributed across the two datasets. As a result, we observe discrepancies when aggregating the total spend received by merchants in some postal sectors in our dataset, and the equivalent spend to merchants in the same postal sector in the merchant dataset. The difference stems from how spending flows are apportioned, and the total expenditure across both datasets in each MCG is the same. We note this additional source of error for any analysis that relies on merchant counts or concentration measures, as well as the reduction in merchant coverage due to statistical disclosure control.

3 Empirical methodology

3.1 Desirable properties of retail market estimates

We identify local markets based on the similarity in card spending flows between postal sectors, in a given month and merchant category group (MCG). The aim is to create clusters of postal areas which we call "local retail markets" that satisfy the following three properties:

1. **Substitutability.** Merchants in the collection of postal sectors must be substitutable for consumers, and vice versa.

2. **Geographical proximity.** Postal sectors in a cluster are geographically close. We expect the majority to be contiguous with other postal sectors in the same cluster.

3. **Appropriate size.** Clusters should match moments of existing local market definitions from the academic literature and competition cases.

Only the first of these properties is imposed through our methodology. By definition, a larger proportion of spending flows from one location to another increases the likelihood that the areas end up in the same cluster. Economically, local residents consider merchants in those postal areas to be substitutable. The reverse is also true: merchants in those areas draw on potential consumers within the same collection of postal sectors. The logic is similar to the CMA's approach in merger cases, which defines local markets as a collection of postcodes that capture 80% of a store's sales or customers (CMA, 2023c).

The second and third assumptions are not imposed on the data and are testable. Geographical proximity and contiguity are readily measurable in the data. To match moments of other market definitions, we have consulted CMA case documents to determine the average distance travelled within the local markets in past cases. We have also obtained estimates on the average distance travelled in retail sectors from the National Travel Survey (DfT, 2023). We use these external estimates to validate the local markets we obtain through the clustering methodology, by setting clustering thresholds that give similar average travel distances.

3.2 The clustering algorithm

This subsection describes the logic of the clustering algorithm we use. The next two subsections describe two key components in more detail: how we compute the dissimilarity score, and how we select the threshold value.

We follow the general approach in Batch et al. (2023) and apply a hierarchical clustering algorithm to determine local markets. However, our approach differs in a few key ways. We calculate a matrix representing the dissimilarity between all cardholder postal sectors $D_{i,j}$ based on spending flows from consumer postal sector i and j to all merchant postal sectors. As an initial condition, we take each postal sector to be its own separate cluster. At each step, the postal sector with the lowest dissimilarity score is added to its nearest cluster. The algorithm recalculates the dissimilarity score between the new cluster and all other clusters. We do this so that *within-cluster* differences between spending flows are small, and *between-cluster* differences are large.

The agglomerative hierarchical clustering algorithm is applied to each pair of clusters, using average linkage. Average linkage ensures that the distance between each cluster is the average distance between every point in one cluster to every point in another cluster (Macklin, 2018), such that

$$D_{a,b} = \frac{1}{N_a * N_b} \sum_{i \in C_a} \sum_{j \in C_b} D_{i,j} \tag{1}$$

where N_a and N_b are the number of postal sectors in clusters A and B, and $\sum_{i \in C_a} \sum_{j \in C_b} D_{i,j}$ is the element-wise sum of the distance matrices for each postal sector in cluster C_a with each postal sector b in cluster C_b. The process is repeated until the threshold H is reached, such that if $D_{a,b} > H$, then A and B do not merge. The resulting clusters contain groups of postal sectors that have the smallest average distance. This distance metric represents how much each cardholder postal sector spends at each merchant location. Hence, the clusters satisfy the property of substitutability.

We apply the algorithm separately to data from each month and MCG. We also recreate these clusters for different threshold levels, but ultimately choose H to match distance related moments described in appendix D.

Intuitively, the method described here differs from the hierarchical clustering algorithm in Batch et al. (2023) in that dissimilarity is measured between consumer postcode sectors over spending flows to all merchant postcode sectors, instead of over total flows between all postcode sectors (implicitly treating inflows and outflows symmetrically). This symmetry assumption may well be

Table 1: Example case of two postal sectors assigned to the same cluster

	CH41 2	L1 1
CH45 5	80.1%	19.9%
CH49 0	80.1%	19.9%

justified when clustering US countries (which contain both residential and retail spaces), but is unsuitable when applied to the smaller geographical areas we study in this article (where many postcode sectors may only contain cardholders, or only contain merchants).

3.3 Measuring dissimilarity

A key component of the methodology is the construction of a dissimilarity measure. This represents how similar two cardholder postal sectors are to one another, based on the merchant postal sectors cardholders in these postal sectors shop at. We use this measure to determine how likely it is that those postal sectors are in the same local retail market.

We begin with a matrix mapping postal sectors into postal sectors. The values are the total spend between each merchant and cardholder location. We measure dissimilarity as cosine dissimilarity. Cosine dissimilarity measures the angle between two vectors, in a multidimensional space, and determines whether the two vectors are approximately facing the same direction (Han et al., 2012). In our data, each vector is a merchant location and each direction is determined by the spend obtained from each cardholder location. Cosine dissimilarity prioritises direction (which location the spend comes from) over the magnitude of spend when calculating dissimilarity.

The equation for the cosine distance between vectors n and m is as follows:

$$1 - \frac{n \cdot m}{\| n \|_2 \| m \|_2} \tag{2}$$

Where $\| * \|_2$ is the 2-norm of its argument *, and $n \cdot m$ is the dot product of n and m.

For a concrete example, take Table 1. The two cardholder postal sectors deemed most similar to one another, according to our dissimilarity score in food and grocery spending during March 2024, are CH45 5 and CH49 0. Each location spends eighty percent of its total expenditure in CH41 2 and the remaining twenty percent at L1 1. This indicates cardholders in these locations shop at a similar set of stores and therefore share a local market. Consequently, the algorithm assigns them to the same cluster.

3.4 Defining thresholds

When defining thresholds, we face a trade-off. On the one hand, we want the majority of consumer spend to fall within markets, to capture the idea that markets define the relevant choice set for consumers. On the other, unrealistically large clusters exaggerate how far people travel and create the illusion that consumers face more choices than they do in practice. We select the cutoff so that the geographical area covered by a local market on average is consistent with external evidence of how far people travel to purchase similar goods. We look at four different metrics to select an appropriate threshold, and compare them to external evidence where possible. These are:

1. The number of local markets in the UK;
2. the average area covered by a market;
3. the average weighted distance travelled per market; and
4. the proportion of spending flows observed within a market.

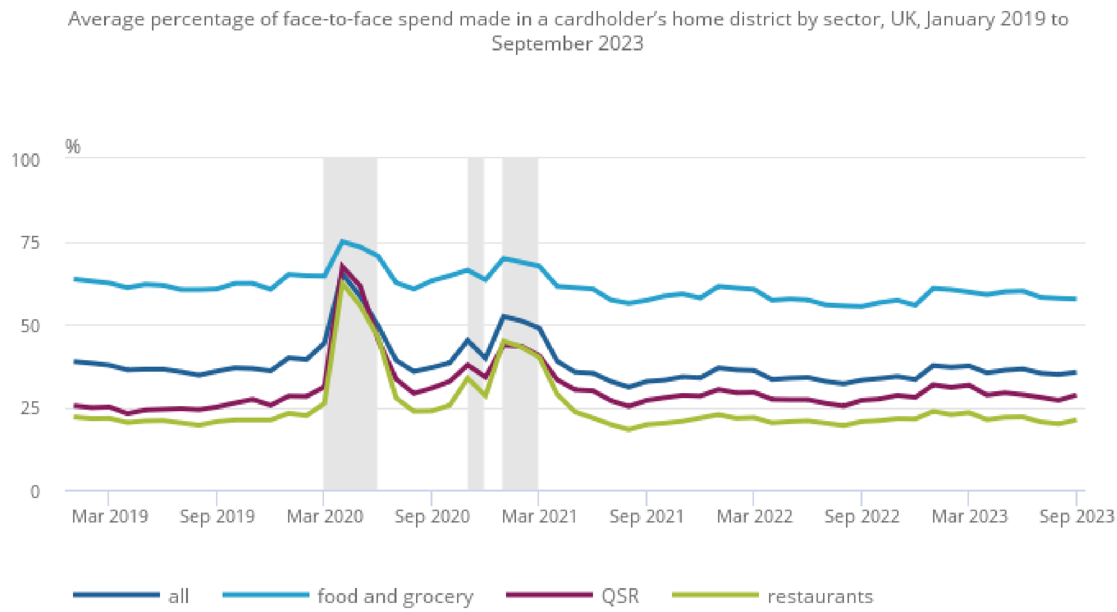
To compute the average area covered by markets, we link our clusters to geospatial information provided by ONS (2016). To calculate the average weighted distance we assign each spending flow to a market based on the consumer location. We then calculate the average postal sector centroid-to-centroid distance across those flows. We note that these distances do not account for variation in population densities *within postal sectors*. For the same reason, distances within the same postal sector are assigned a value of 0 in the distance calculations. As a result, the distance estimates are likely an underestimate.[8]

We weigh distances according to the proportion of spend that occurs within a cluster. Then we take the average of this across all clusters. To give an example, suppose a cluster consists of just one postal sector, postal sector A. We return to the spending flow data, see that the total spend from consumers in postal sector A to merchants in postal sector A was £70, the total spend from consumers in postal sector A to merchants in postal sector B was £30, which is 2 miles from postal sector A. We calculate the weights such that spend within postal sector A is weighted $0.7 = (70/(70+30))$. We calculate the weighted distance of the cluster as $0.7 \cdot 0 + 0.3 \cdot 2 = 0.6$ miles.

Finally, we calculate the proportion of spending flows occurring within a cluster. In the example above, the cluster consists of one postal sector, postal sector A. £70 is spent within that cluster,

[8] To deal with this downward bias, one could randomly draw postcodes from each postal sector to represent either the business or the consumer or both.

Figure 3: Proportion of spend occurring within the same postal district, over time, on average for food and grocery, restaurants and quick-service restaurants (QSR). Source: ONS analysis, ONS (2024b)



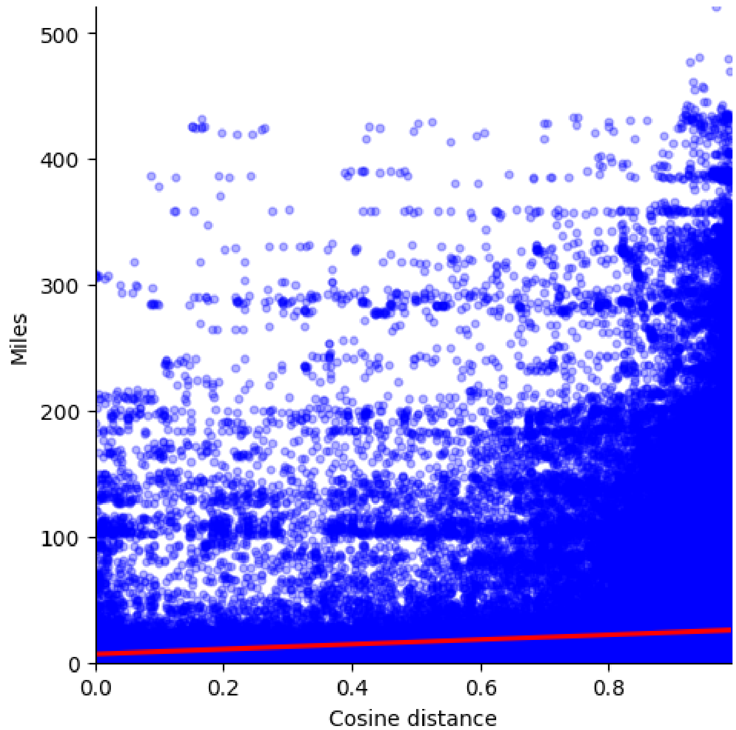
and £30 is spent to merchants in postal sector B, which is outside the cluster.

3.5 Descriptive statistics on card spending flows

This section presents descriptive statistics on card spending flows over time and the distribution of cosine dissimilarity scores to help assess the accuracy of our approach. For interested readers, ONS (2023a) and ONS (2023b) go into much further detail regarding patterns in the Visa card spending data. Here, we focus on how much spending occurs within small geographical areas. This can be useful to gauge whether our data is of sufficient geographical granularity, and to interpret the weighted distances we calculate. Figure 3 shows that in some MCGs, like food and grocery, close to 60% of expenditure occurs within the the same postal district. For others, such as restaurants, the proportion is instead closer to 20%. For most MCGs, we observe spikes in the proportion of own-postcode district spending during periods of pandemic-related national mobility restrictions.

Figure 4 shows the relationship between dissimilarity scores and the geographic distance between cardholder postal sectors. Each point represents the cosine dissimilarity between each pairwise combination of postal sectors, and the distance in miles between those postal sectors. The red line

Figure 4: Geographic distance against cosine dissimilarity for postal sectors in food and grocery, UK, June 2023

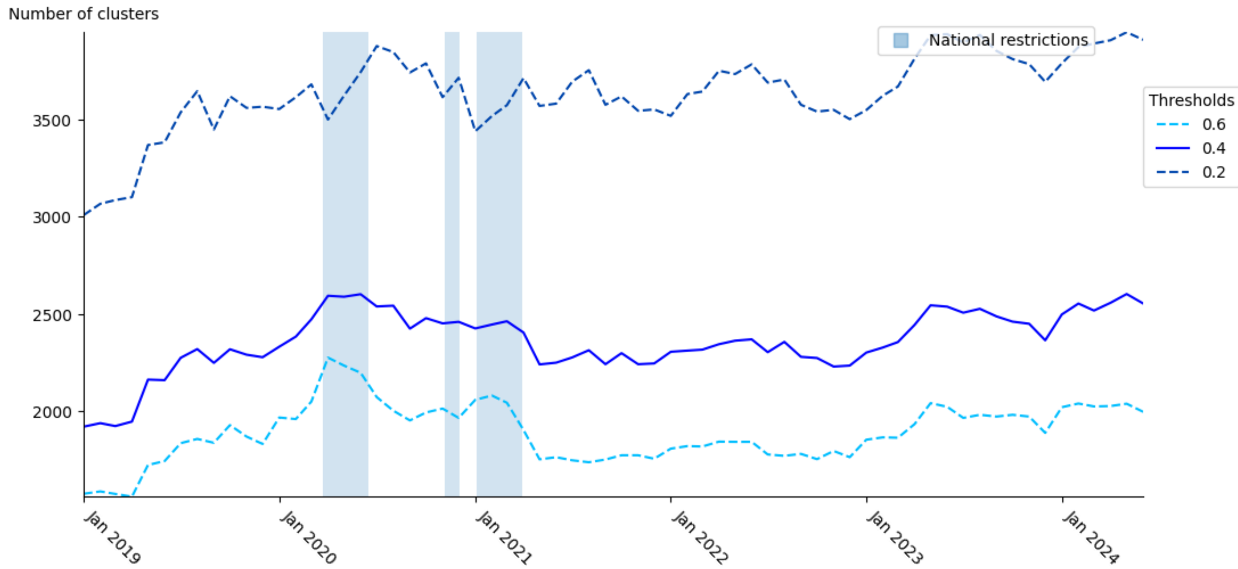


indicates a linear line of best fit. As expected, the geographic distance increases with dissimilarity. This indicates that most cardholder locations in the UK do not spend equally across merchant locations as all other cardholder locations. Only cardholder locations geographically close to each other spend at similar locations, and would therefore be expected to have a low dissimilarity score. These combinations of locations make up a small fraction of the data.

4 Estimates of UK local retail markets

This section presents our main results. We first show how the number of local retail markets, average market size, travel distance and within-market spending has changed over time, for our baseline threshold choice. We focus on food and groceries for illustrative purposes, but the appendix reports results for other sectors and threshold choices. We then map out local retail markets for a selection of cities and towns, and a selection of MCGs, as well as the UK as a whole. We analyse the degree of spatial correlation across MCGs. Finally, we provide demographic correlates of local market size.

Figure 5: Number of local retail markets by threshold value, UK, Food and Grocery



4.1 The number of UK local retail markets

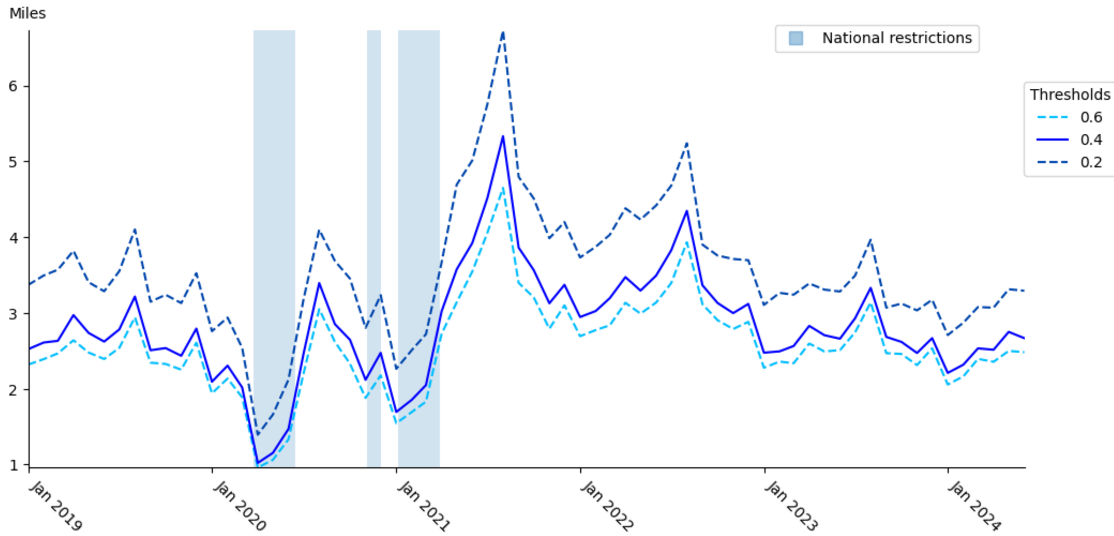
Figure 5 presents the number of local food and grocery retail markets for three different values of the threshold, on a monthly basis between January 2019 and June 2024. The monthly count of local retail markets in food and grocery using the central 0.4 threshold is plotted in dark blue, with alternative 0.2 and 0.6 thresholds in dotted light blues. The periods of UK national restrictions are shaded in for reference. There is a clear change in the measure during lockdown periods, as people adapted to mobility restrictions to change their spending patterns.

As our cleaned dataset contains just over 11,000 postal sectors in the UK, an average of about 2,500 local retail markets for our baseline threshold value of 0.4 implies that on average a local retail market in food and grocery contains about four to five postal sectors. Figure B4 in the appendix shows the number of local retail markets in each UK region in our dataset. The coverage of Northern Ireland is limited in our data, so special care has to be taken when interpreting results for Northern Ireland.

4.2 Market size and travel distance

Figure 6 shows the median weighted average travel distance per cluster, in miles, in the UK. This is based on ONS estimates of the centroid-to-centroid distance between postal sectors, which we combine with our spending flows. For food and groceries, the median travel distance is roughly two

Figure 6: Median cluster size, UK, food and grocery



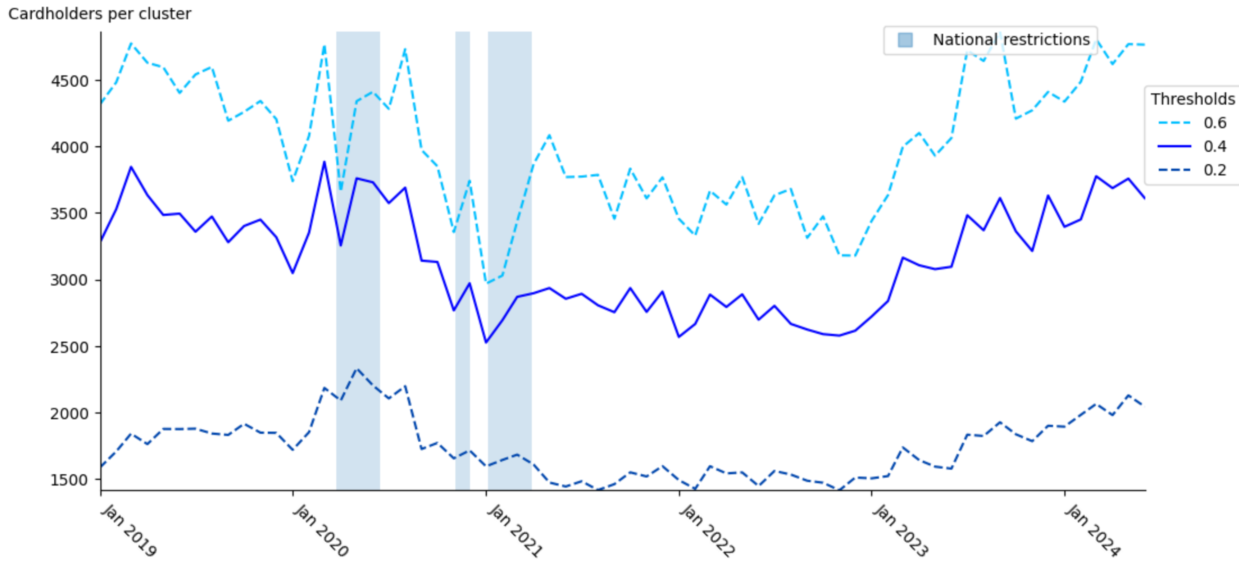
to three miles. Figure B6 in the appendix breaks the UK average down by region. At one end of the spectrum, local retail markets in London are about half the size of the UK average. At the other, local markets in the South West are about twice the size of the UK average.

Figure B9 in the appendix shows the proportion of UK local markets containing a single postal sector, when applying different clustering thresholds. For our baseline threshold level, about 40% of food and grocery markets consist of single postcode sectors. This reflects the small market size for this particular MCG. However, travel distances are not the only reason why local markets consist of single postcodes. Where data is patchy for coverage or disclosure reasons, postal sectors may not closely match any other postal sectors in the dataset and therefore not be assigned to nearby clusters.

4.3 Within-market spending

When considering measures such as spend, transactions, and cardholders per market, a few additional caveats apply. These measures are not solely related to geographic area size, but also to data coverage and alternative consumer purchase channels. If consumers spend more heavily online, then the coverage of face-to-face transactions in our data diminishes. This means that the total spend per cluster will decrease, even if people still travel the same distance, and the geographic shape of the true underlying market remains the same. Additionally, when people change how far they are willing to travel, the changing composition of the markets may also trigger changes in the total

Figure 7: Total number of cardholders per cluster, UK, food and grocery



amount of spending, the number of transactions, and the number of cardholders within a given market.

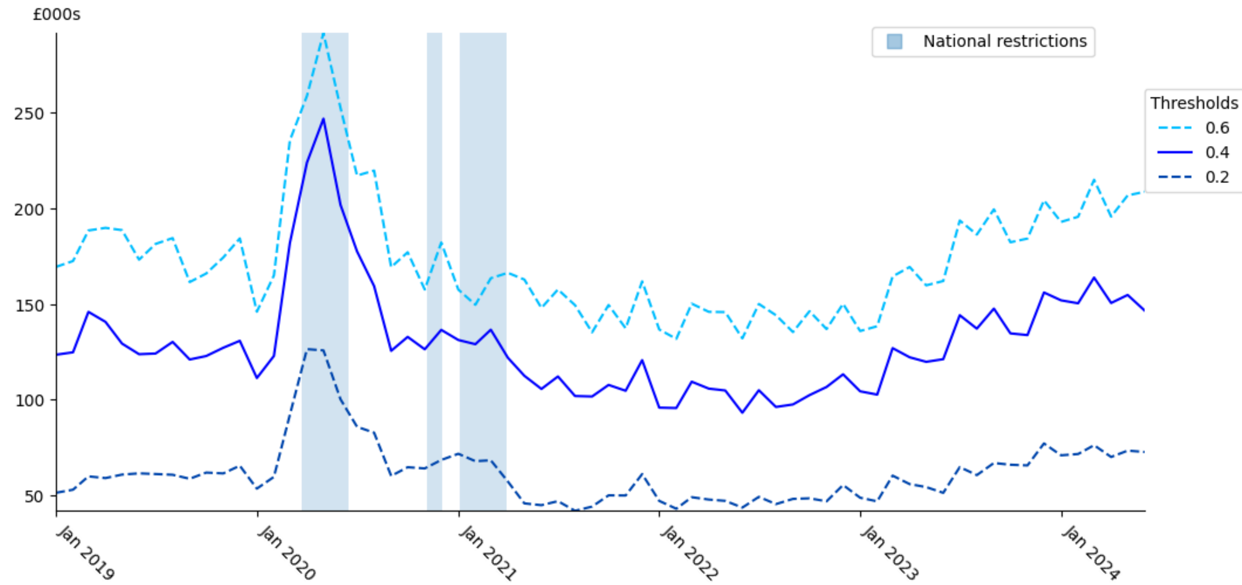
Figure 7 shows the average number of cardholders within a local market for food and groceries, for different thresholds. On average, for our baseline threshold a local market contains about 3,500 cardholders. This number has fallen somewhat in the pandemic as markets have shrunk, and expanded again more recently. Figure B7 in the appendix breaks the total amount of cardholders within a cluster, for the different UK regions.

Figure 8 shows the total amount spent within the average market, for three different threshold values. The average monthly card spend hovers around £150,000, with the exception of a sharp, short spike during the first national Covid-19 lockdown, perhaps reflecting widely reported panic-buying. Total spend levels do not seem sensitive to the threshold level either. Figure B8 in the appendix breaks this average down into the UK regions. Average monthly spend per market in London is almost double the national average, with the South West and North East closer to half the national average.

4.4 Mapping GB retail markets across space

Figure 9 shows the the geographic mapping of local retail markets for food and groceries across GB in March 2024. Northern Ireland, while included in the dataset and estimation, is omitted from the maps due difficulties accessing NI-specific within the ONS' cloud-computing platform. The map

Figure 8: Total amount spent per cluster, UK, food and grocery



represents the cardholder postal sectors assigned to each cluster group. Clusters are coloured by merchant counts: the darker the colour, the more merchants in a local market. This merchant count provides a first proxy measure of local market concentration.[9] Due to the high amount of clusters across GB, some adjacent clusters use the same colours. However, the geographic continuity property is still clearly visible.[10] Postal sectors are shaded in with a diagonal pattern denote missing data. Figures B1 to B3 in the appendix show the spatial distribution of retail markets for the remaining MCGs.

In addition to these GB-wide maps, for illustrative purposes we also zoom into a few representative cities and towns across the UK. Panel (a) of Figure 10 shows the area surrounding Darlington in County Durham, a market town where the CMA's Microeconomics Unit is based. The map shows smaller local markets concentrated on major local towns, such as Northallerton and Newton Aycliffe, and larger retail markets spanning the rural communities, for instance in the North Pennines. Edinburgh, Cardiff and Manchester show similar patterns with smaller markets in more densely populated areas.

[9]We also compute bounds on a local retail markets Herfindahl-Hirschman Index (HHIs). What makes computing HHIs in this context difficult is that we do not observe the distribution of spending flows across merchants *within the same merchant postcode*. We therefore compute initial bounds in the following way. The upper bound is computed by assigning all spend flowing to a merchant postal sector to a single merchant. The lower bound is computed by assigning equal shares of the total inbound spending flows to each market. The bounds in the current version of the dataset are so wide as to be relatively useless in practice, and are therefore not reported in this draft.

[10]Alongside this draft, we have released a dataset of MCG-by-month local market estimates. We hope this allows researchers to visualise and use these estimates according to their own analytical needs.

Figure 9: Cluster locations across Great Britain, food and groceries, March 2024

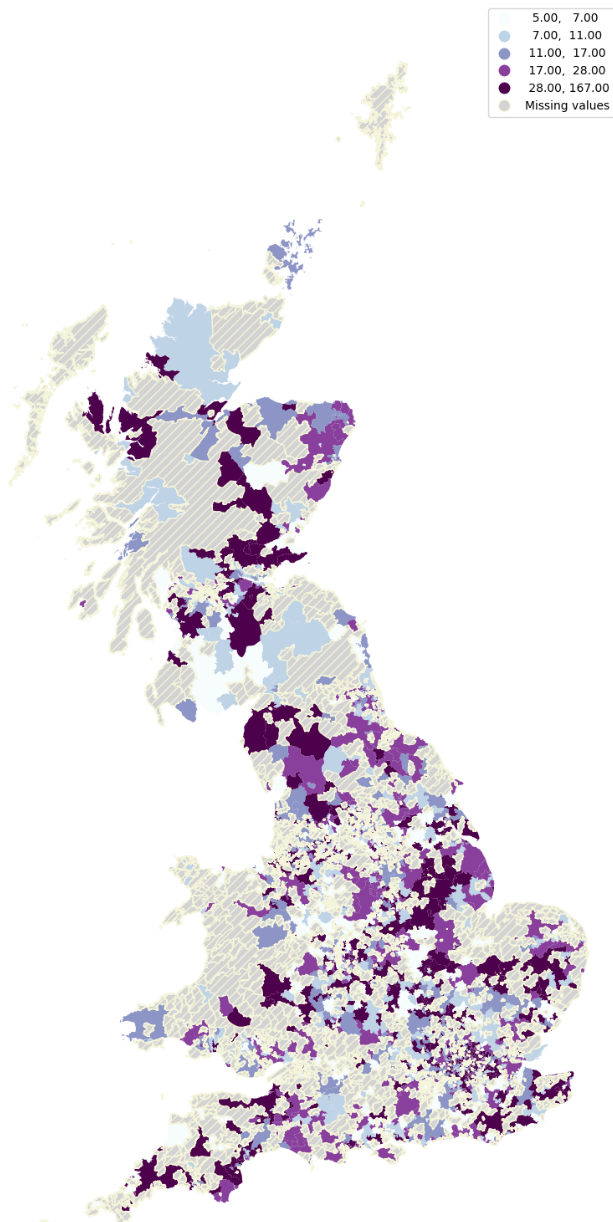
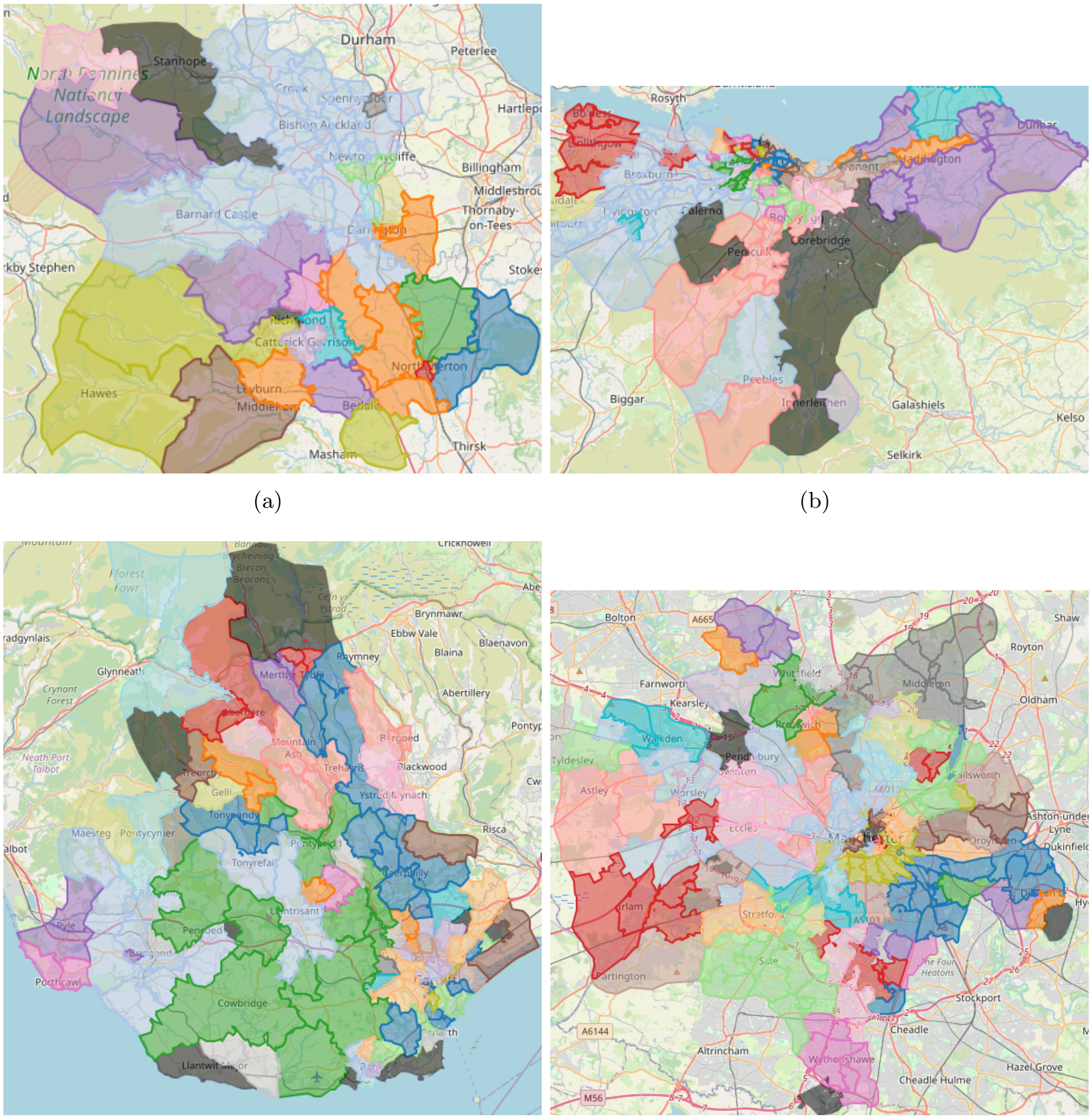


Figure 10: Cluster locations across select UK cities, food and groceries March 2024: (a) Darlington (b) Edinburgh (c) Cardiff (d) Manchester

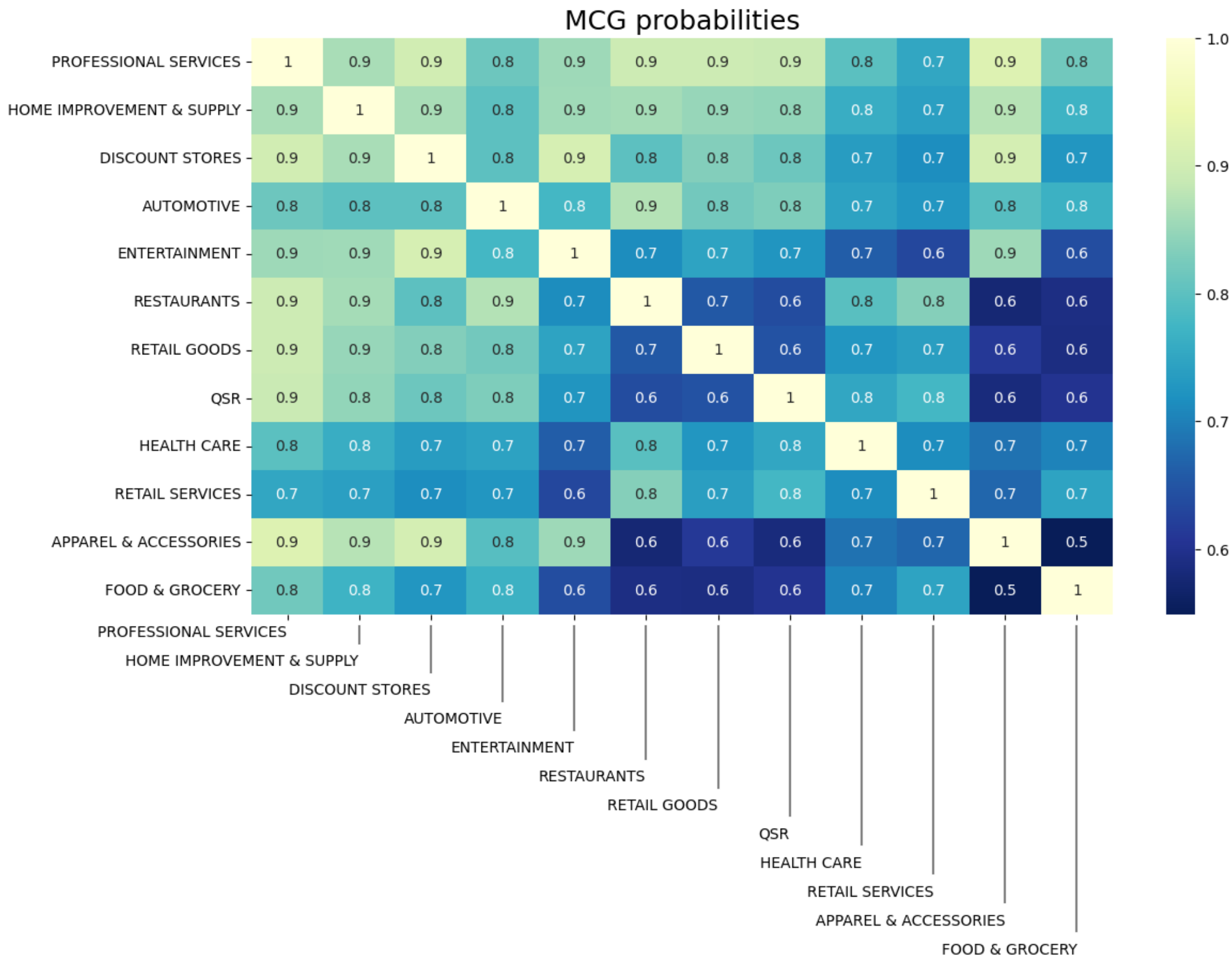


(a) (b) (c) (d)

4.5 Cross-market and spatial correlations in market boundaries

Correlations of market definitions across MCGs can provide information about consumer behaviour patterns. For instance, in a world where suburban households drive to the same shopping malls for all their retail and hospitality, card spending flows will display high correlations across MCGs. Conversely, if households shop for groceries and restaurants locally and other, more expensive consumer items more widely, cross-MCG correlations will be lower. Figure 11 shows the spatial

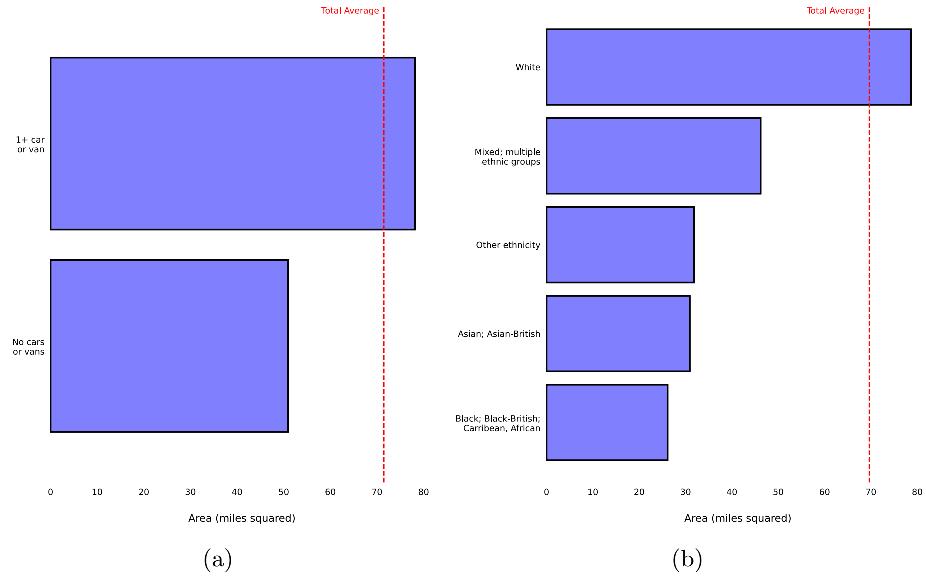
Figure 11: Average market overlap across merchant good categories (MCGs)



correlations across MCGs. We obtain these correlations by computing the probability that two postal sectors are in the same local market for MCG j, conditional on being in the same market for MCG i. We average these probabilities over all postal sectors. MCGs for which less than 2,000 out of approximately 12,000 postal sectors remain after disclosure and dominance controls are dropped from the table.

Figure 11 shows that local markets for home improvement and supply stores, discount stores, automotive retailers, entertainment and restaurants are relatively highly correlated with other MCGs. By contrast, healthcare, apparel and food and groceries display relatively low correlations with other MCGs. This may reflect differences in market scale, but further work is needed to understand these correlations.

Figure 12: Average market size (in square miles miles), in food and groceries, based on demographic characteristics around: (a) car/van availability (b) ethnic group



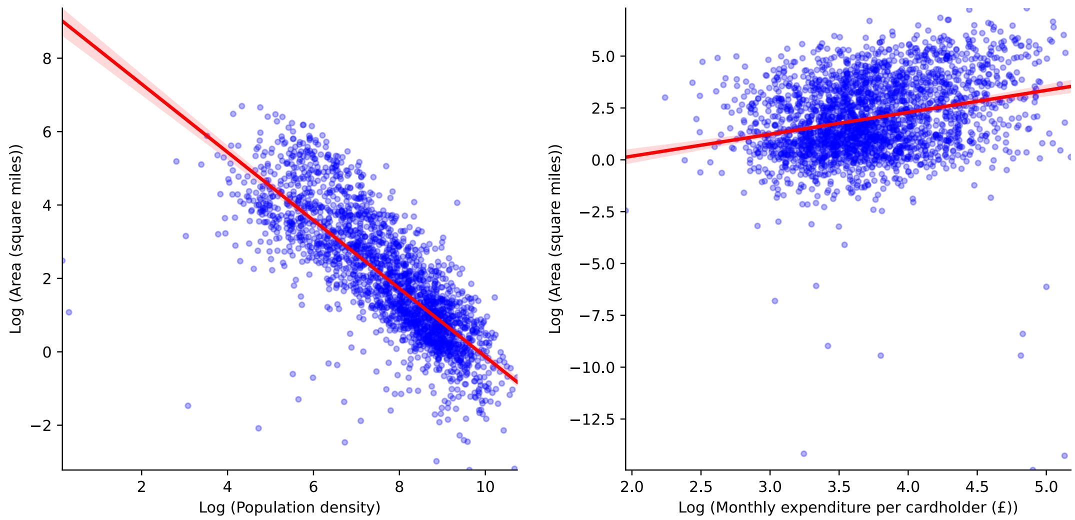
4.6 Market size and demographics

Using ONS census data ONS (2024a), we calculate the average size of a market based on demographic characteristics in the component postal sectors. In this subsection, we show correlations between local market size and the following demographic attributes: car/van availability for residents, their ethnic group, gender, the household composition, age, and accommodation type. Local market size is not systematically related to gender, but shows clear relationships to other demographic characteristics. As Figure 12 and Figure C2 in the appendix show, local markets are larger in areas with heavy car ownership, with a larger white population share, with single-family households, older residents and more rural or suburban forms of dwellings (detached, semi-detached and mobile housing).

Appendix C shows regression results when we control for these characteristics separately and jointly, for four example MCGs. Appendix section C.1 shows how these estimates vary over time. These time-varying estimates yield rich information how consumer search and travel patterns change over time as a function of demographic and economic characteristics. A full exploration of these patterns is left for future work.

Finally, Figure 13 shows the distribution of local market size for food and groceries against population density (Panel A) and average monthly cardholder expenditure (Panel B). Local market

Figure 13: Average market size (square miles) against population density and average expenditure per cardholder



size is decreasing in population density, and increasing in cardholder expenditure, a measure of local incomes. These differences likely reflect both demand-side (consumer density and willingness-to-pay) and supply-side (economies of scale) factors.

4.7 Robustness to threshold selection

Our time trends are not particularly sensitive to the exact choice of threshold, though of course market size estimates can vary considerably. Appendix D shows a range of robustness checks for different threshold levels. For example, our baseline 0.4 threshold yields an average local retail market size of about 24 square miles in outside the pandemic period, but this increases to just over 33 miles at a 0.5 threshold and falls to close 16 miles at a 0.3 threshold. The appendix tables similarly show, for each MCG, how the number of local retail markets, average travel distance, local market area, within-market spend, within-market transactions and other characteristics change across all plausible threshold values. We show these statistics separately for the Covid-19 pandemic period, and the non-pandemic baseline, as well as for all MCGs. We hope this allows researchers to understand the importance of threshold selection, and to make their own choices when applying this methodology going forward.

5 Conclusion

Defining local markets is crucial for our understanding of competition, from both a research and competition enforcement point of view. But good data on local spending patterns is expensive and hard to come by. In this paper, we use monthly, spatially highly disaggregated payment card spending data for twenty-three different retail merchant category groups to construct geographical retail markets for the UK. We cluster cardholder postal sectors based on the similarity of their spending flows across merchants for a given retail category. We choose category-specific thresholds for the clustering algorithm to match average travel distances for similar products from existing survey data.

As a first proof of concept, we apply this technique to estimate local retail markets for food and groceries. We obtain roughly 2,000 local food retail markets. Estimates generally are qualitatively and quantitatively similar to existing studies, but market size varies considerably across the UK. While we do not impose geographical contiguity, most local markets do in fact consist of adjacent postal areas. The analysis in this paper allows us to better understand the size, shape and persistence of local markets, the correlation across different merchant categories and the demographic, economic and geographical determinants of local markets. Focusing on a few example towns and cities, we find that, unsurprisingly, local markets are smaller in densely-populated urban areas. Geographically, London is an outlier in both its high amount of spending per market and its small market size. Merchant category groups (MCGs) also differ systematically in the degree of spatial correlation across their markets, with local markets for home improvement stores for instance much more closely correlated with other MCGs than for instance food and grocery markets.

We then characterise geographical, economic and demographic determinants of market size. Local market size is systematically related to consumer characteristics such as age, dwelling type, mode of transport and ethnic group. Finally, we estimate merchant counts in each local market as a first local retail market concentration measure. Alongside the paper, we will make our retail market estimates available to the research community. We hope these new estimates will enable researchers to better understand local competition, firm dynamics and consumer behaviour.

References

E. Audy and C. Erutku. Price tests to define markets: An application to wholesale gasoline in canda. *Journal of Industry, Competition and Trade*, 5:137–154, 2005.

D. Autor, C. Patterson, and J. Van Reenen. Local and national concentration trends in jobs and sales: The role of structural transformation. Technical report, National Bureau of Economic Research, 2023.

A. Batch, B. Bridgman, A. Dunn, and M. Gholizadeh. Consumption zones. 2023.

C. L. Benkard, A. Yurukoglu, and A. L. Zhang. Concentration in product markets. Technical report, National Bureau of Economic Research, 2021.

D. P. Byrne, N. de Roos, M. S. Lewis, L. M. Marx, and X. A. Wu. Asymmetric information sharing in oligopoly: A natural experiment in retail gasoline. *Available at SSRN 4513435*, 2024.

S. Calligaris, M. Chaves, C. Criscuolo, J. De Lyon, A. Greppi, and O. Pallanch. Exploring the evolution and the state of competition in the eu, 2024.

CMA. Sainsbury's, asda, phase 2 final report. https://assets.publishing.service.gov.uk/media/5cc1ec1340f0b64031cfa6f0/Final_reportSA.pdf, 2017. [Accessed 08-01-2024].

CMA. Competition, choice and rising prices in groceries. https://assets.publishing.service.gov.uk/media/64a280e845b6a200123d46e7/Supply_of_road_fuel_in_the_United_Kingdom_market_study_Final_Report.pdf, 2023a.

CMA. Competition, choice and rising prices in groceries. https://www.gov.uk/government/publications/competition-choice-and-rising-prices-in-groceries, 2023b. [Accessed 08-01-2024].

CMA. https://assets.publishing.service.gov.uk/media/5a81e8e840f0b62302699d23/retail-mergers-commentary.pdf, 2023c. [Accessed 19-12-2023].

M. Covarrubias, G. Gutiérrez, and T. Philippon. From good to bad concentration? us industries over the past 30 years. *NBER Macroeconomics Annual*, 34(1):1–46, 2020.

N. Crouzet and J. C. Eberly. Understanding weak capital investment: The role of market concentration and intangibles. Technical report, National Bureau of Economic Research, 2019.

J. De Loecker, J. Eeckhout, and G. Unger. The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, 135(2):561–644, 2020.

DfT. Purpose of travel — gov.uk. https://www.gov.uk/government/statistical-data-sets/nts04-purpose-of-trips, 2023. [Accessed 19-12-2023].

A. Eizenberg, S. Lach, and M. Oren-Yiftach. Retail prices in a city. *American Economic Journal: Economic Policy*, 13(2):175–206, 2021.

K. G. Elzinga and T. F. Hogarty. The problem of geographic market delineation in antimerger suits. *The Antitrust Bulletin*, 18(1):45–81, 1973.

K. G. Elzinga and A. W. Swisher. Limits of the elzinga–hogarty test in hospital mergers: The evanston case. *International Journal of the Economics of Business*, 18(1):133–146, 2011.

C. N. Gibbs, B. Guttman-Kenney, D. Lee, S. T. Nelson, W. H. van der Klaauw, and J. Wang. Consumer credit reporting data. Technical report, National Bureau of Economic Research, 2024.

J. Han, M. Kamber, and J. Pei. 2 - getting to know your data. In J. Han, M. Kamber, and J. Pei, editors, *Data Mining (Third Edition)*, The Morgan Kaufmann Series in Data Management Systems, pages 39–82. Morgan Kaufmann, Boston, third edition edition, 2012. ISBN 978-0-12-381479-1. doi: https://doi.org/10.1016/B978-0-12-381479-1.00002-2. URL https://www.sciencedirect.com/science/article/pii/B9780123814791000022.

C.-T. Hsieh and E. Rossi-Hansberg. The industrial revolution in services. *Journal of Political Economy Macroeconomics*, 1(1):3–42, 2023.

S. Y. Kwon, Y. Ma, and K. Zimmermann. 100 years of rising corporate concentration. *American Economic Review*, 114(7):2111–2140, 2024.

S. Lui, L. Romaniuk, J. Schneebacher, and N. Wolf. The uk longitudinal business database. *ESCoE Discussion Paper*, 2023.

C. Macklin. Hierarchical agglomerative clustering algorithm example in python. https://towardsdatascience.com/machine-learning-algorithms-part-12-hierarchical-agglomerative-cluste 2018. [Accessed 09-01-2024].

S. Martin. *Market transparency and consumer search-Evidence from the German retail gasoline market*. Number 350. Dice Discussion Paper, 2020.

ONS. National Statistics Postcode Lookup. https://www.ons.gov.uk/methodology/geography/geographicalproducts/postcodeproducts, 2016. [Accessed 15-01-2024].

ONS. Regional consumer card spending, uk: 2019 to 2023. https://www.ons.gov.uk/economy/economicoutputandproductivity/output/articles/regionalconsumercardspendinguk/2019to2023, 2023a. [Accessed 16-01-2024].

ONS. Regional consumer card spending trends qmi. https://www.ons.gov.uk/economy/economicoutputandproductivity/output/methodologies/regionalconsumercardspendingtrendsqmi, 2023b. [Accessed 16-01-2024].

ONS. Postal geographies - Office for National Statistics — ons.gov.uk. https://www.ons.gov.uk/methodology/geography/ukgeographies/postalgeography, 2023c. [Accessed 07-05-2024].

ONS. Uk census. https://www.ons.gov.uk/census, 2024a. [Accessed 08-01-2024].

ONS. Consumer card spending, flow of spending across the uk: 2019 to 2023. https://www.ons.gov.uk/economy/economicoutputandproductivity/output/articles/consumercardspendingflowofspendingacrosstheuk/2019to2023#spend-within-a-cardholder-s-home-postal-district-or-postal-area, 2024b. [Accessed 09-09-2024].

C. Patterson and N. J. Vavra. Concentrating on customers: Spending across firms and space. Technical report, 2024.

PSR. Snapshot of payments in the uk over time. https://www.psr.org.uk/media/20ob5wee/payments-over-time.pdf, 2021. [Accessed 08-01-2024].

E. Rossi-Hansberg, P.-D. Sarte, and N. Trachter. Diverging trends in national and local concentration. *NBER Macroeconomics Annual*, 35(1):115–150, 2021.

C. Shapiro and A. Yurukoglu. Trends in competition in the united states: What does the evidence show? Technical report, National Bureau of Economic Research, 2024.

C. M. Tolbert and M. Sizer. Us commuting zones and labor market areas: A 1990 update. 1996.

S. W. Ulrick, S. B. Sacher, P. R. Zimmerman, and J. M. Yun. Defining geographic markets with willingness-to-travel circles. *Supreme Court Economic Review*, 28(1):241–284, 2020.

L. J. White and J. Yang. What has been happening to aggregate concentration in the us economy in the twenty-first century? *Contemporary Economic Policy*, 38(3):483–495, 2020.

A Data description

Our data includes the following Merchant Category Groups (MCGs). Each MCG in turn consists of several product categories. We provide only a sample of some product categories associated with a sample of MCGs. Please note this list is non-comprehensive for disclosure reasons. Our aim is to reduce ambiguity around industry definitions regarding the estimated local markets.

- **Apparel and accessories**: E.g. Clothing Stores
- **Retail goods**: E.g. Books. movies, music, and jewellery stores
- **Food & grocery**: E.g. Bakeries, Supermarkets
- **Fuel**: E.g. Service stations, fuel dealers
- **Healthcare**: Includes health care and veterinary Services
- **Home improvement and supply**: E.g. Carpentry, garden supplies and roofing supplies
- **Restaurants**: E.g. caterers and restaurants
- **Retail services**: Includes for example beauty salons, barber shops, dry cleaners and shoe repair shops.

B Additional figures and tables

Figure B1: Cluster locations across Great Britain, Retail Goods, March 2024

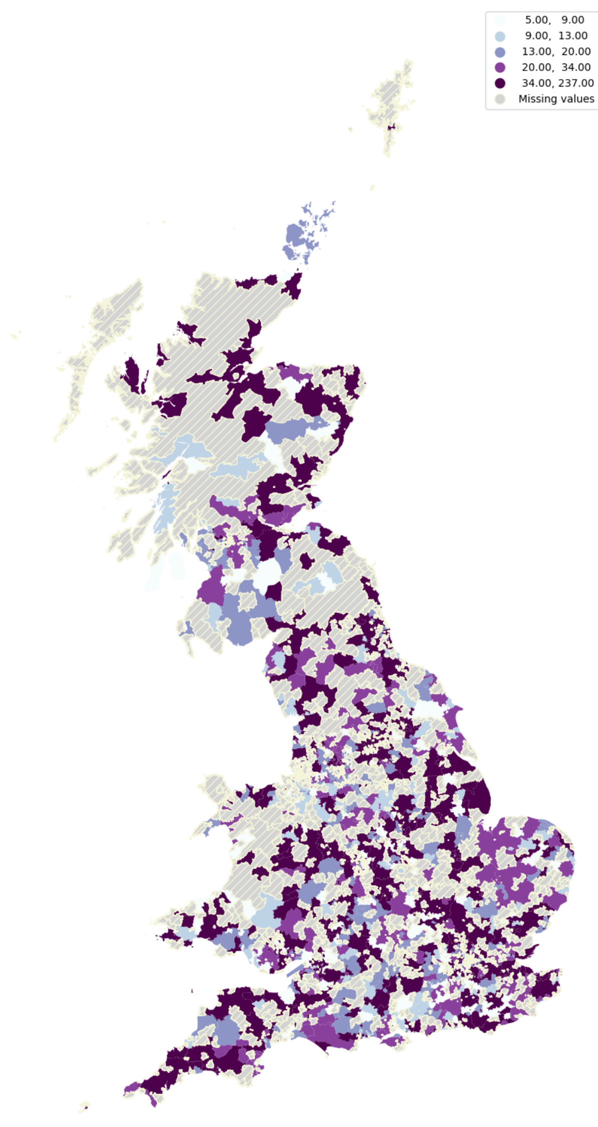
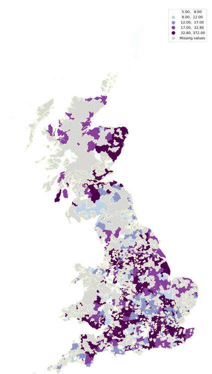
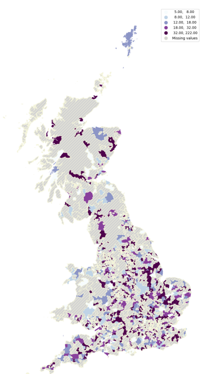


Figure B2: Cluster locations across Great Britain, select merchant category groups (MCGs): (a) Apparel (b) Automotive (c) Business-to-Business (d) Department stores (e) Discount stores (f) Electronics (g) Entertainment (h) Fuel (i) Health Care

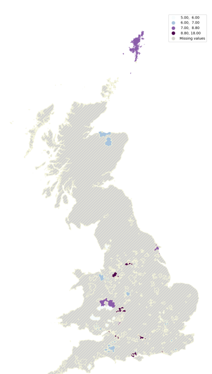


(a)

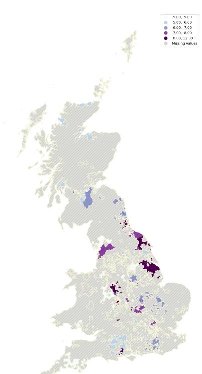


(b)

(c)

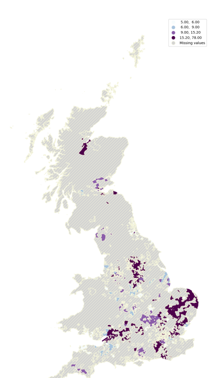


(d)

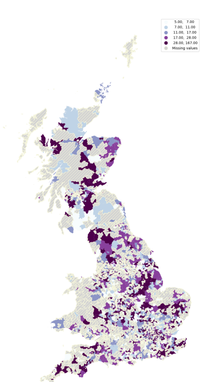


(e)

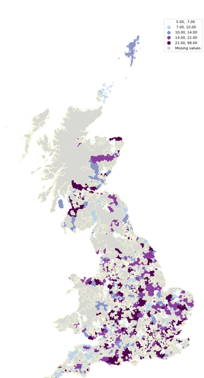
(f)



(g)



(h)



(i)

Figure B3: Cluster locations across Great Britain, select merchant category groups (MCGs): (a) Home improvement (b) Lodging (c) Professional Services (d) Quick Service Restaurants (e) Restaurants (f) Retail Goods (g) Retail Services (h) Telecoms and Utilities (i) Transportation

(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure B4: Number of markets, across regions, monthly, UK, Food and Grocery

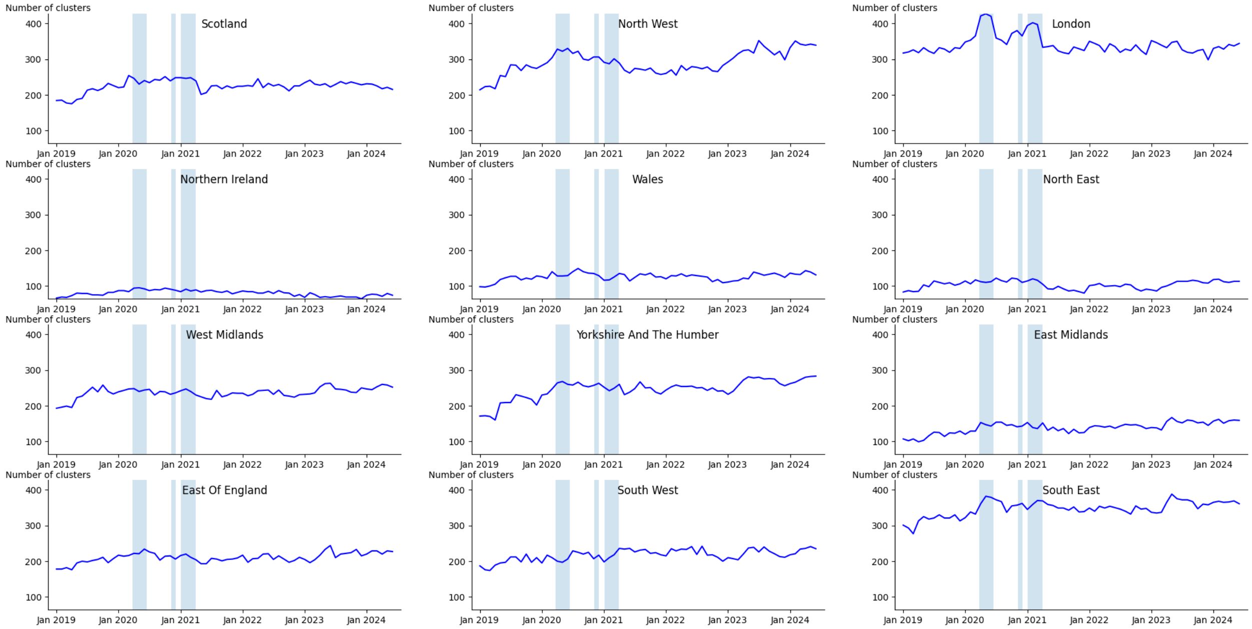


Figure B5: Median market geographic area size (square miles), across regions, monthly, UK, food and grocery

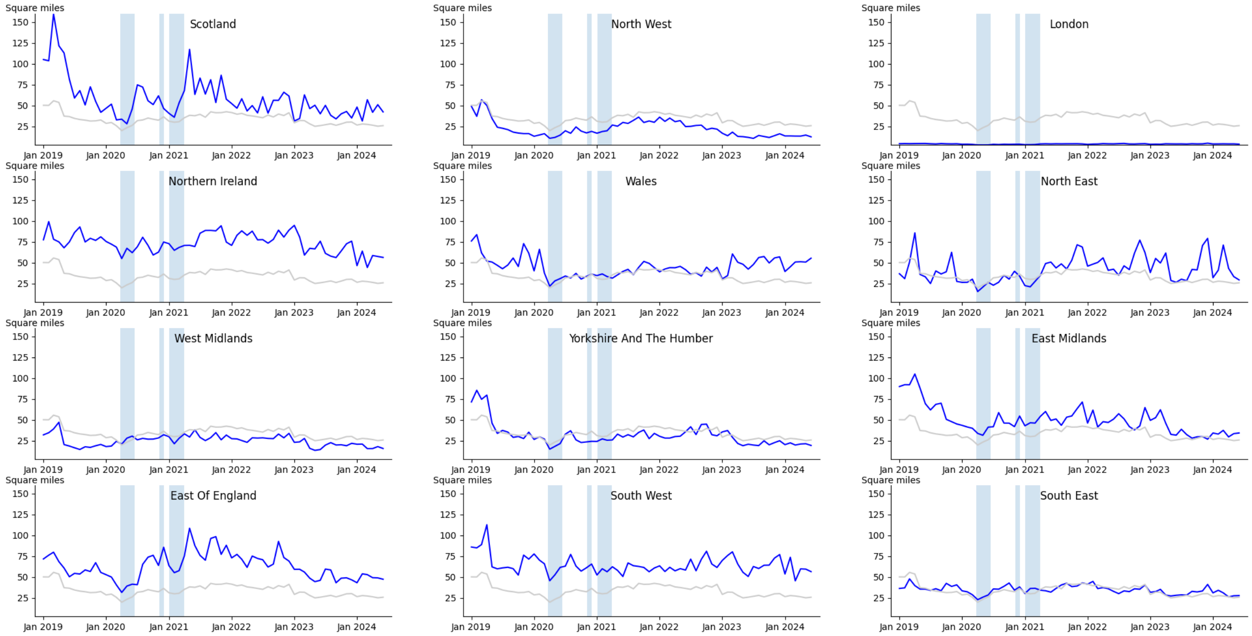


Figure B6: Median weighted (by spend) distance travelled per market (miles), across regions, UK, food and grocery

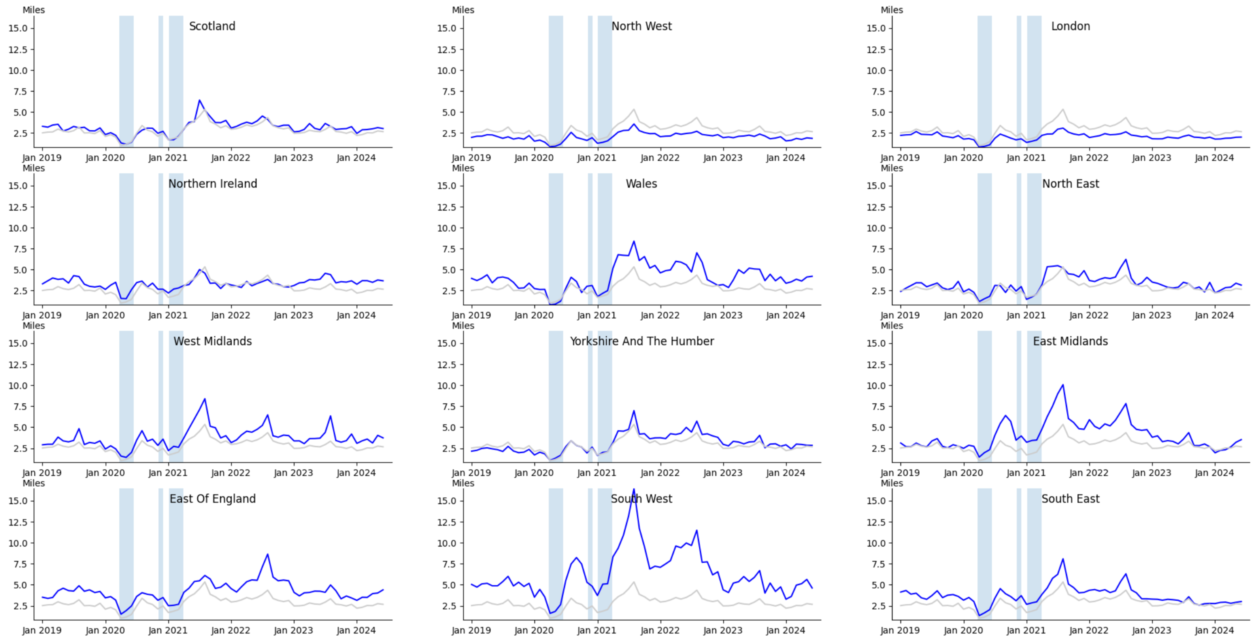


Figure B7: Median number of cardholders per market, across regions, UK, food and grocery

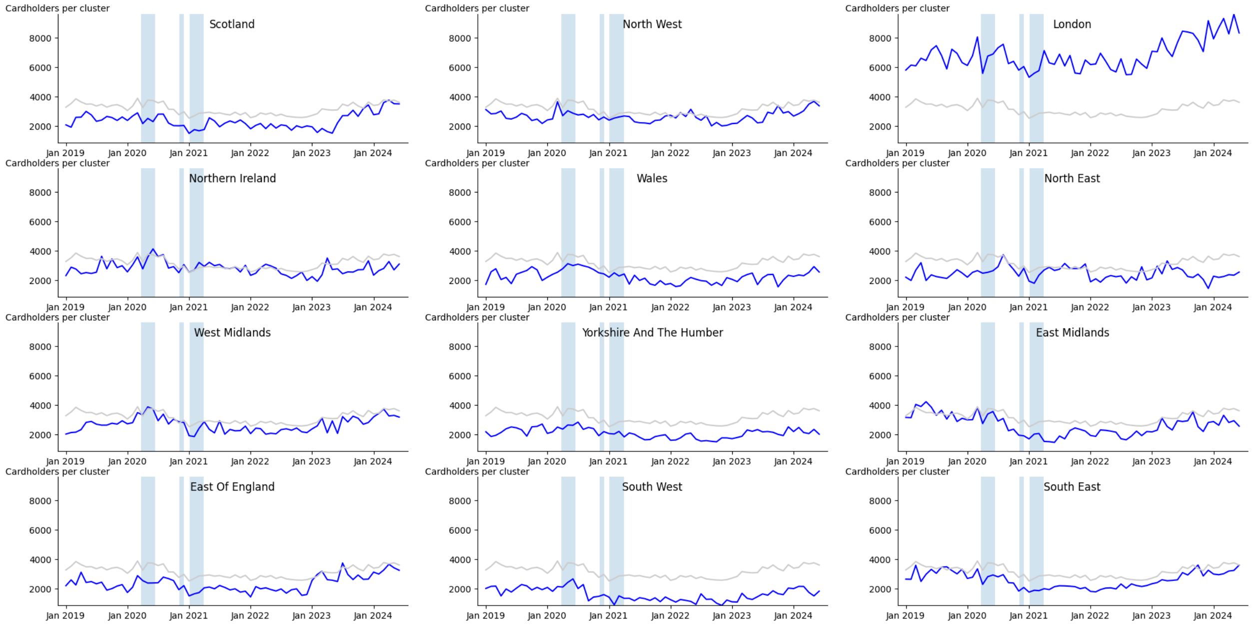


Figure B8: Median expenditure per market, across regions, UK, food and grocery

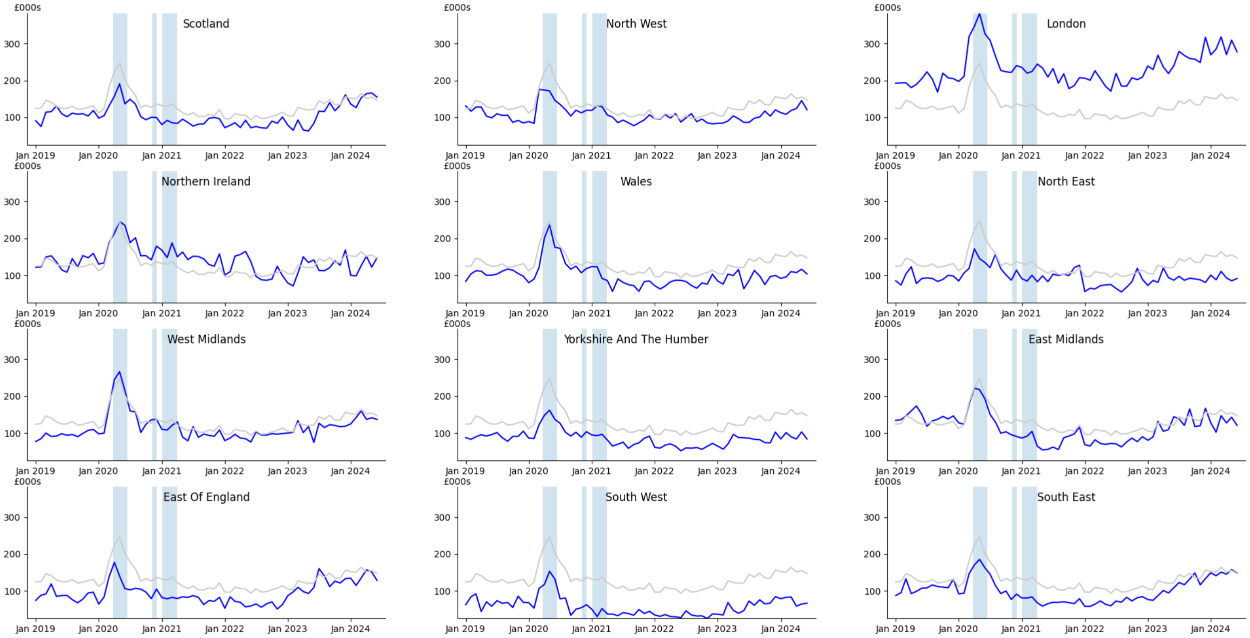


Figure B9: Proportion of clusters based on a singular postal sector, UK, food and grocery

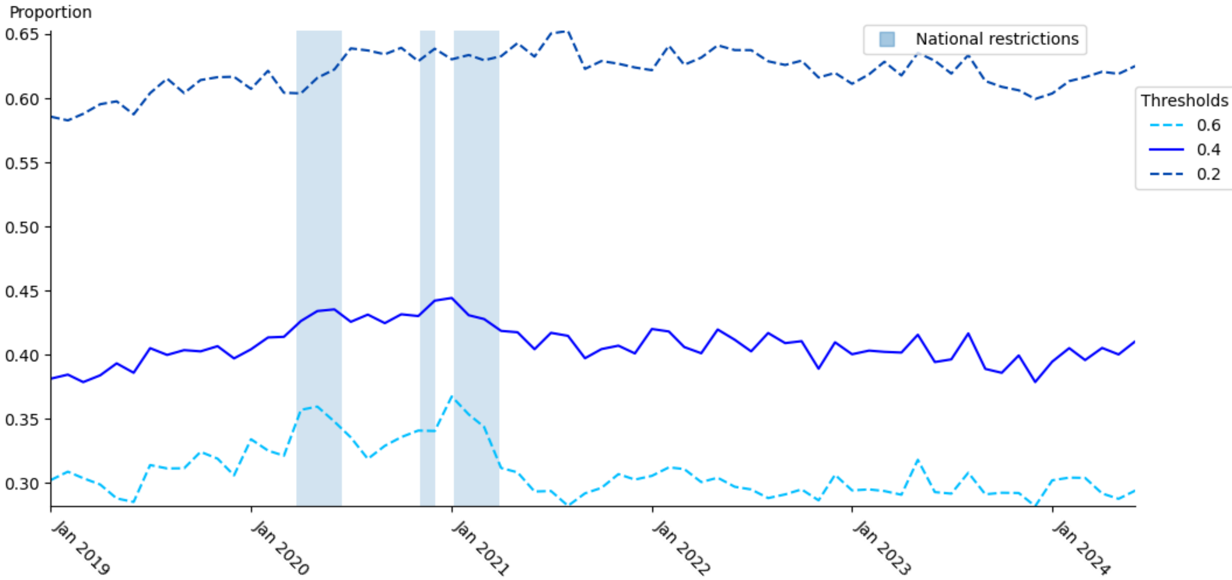
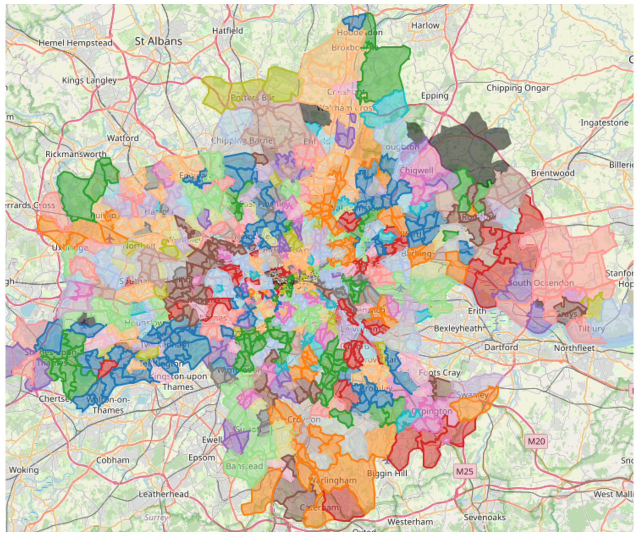
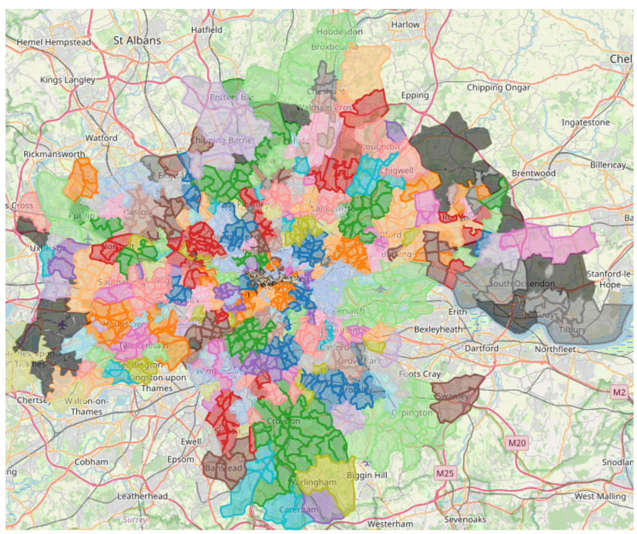


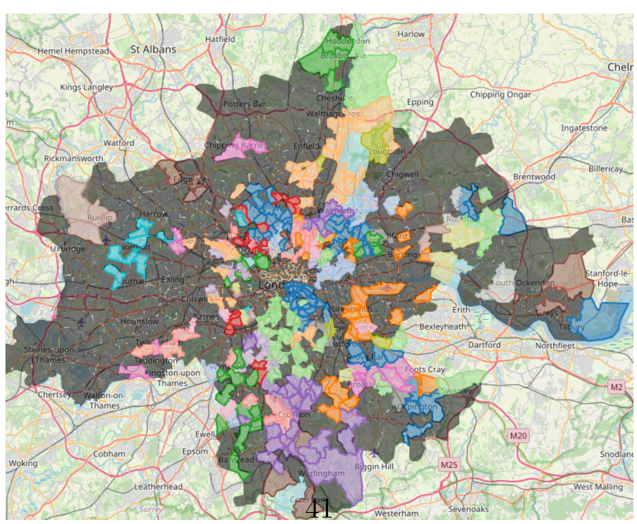
Figure B10: London cluster locations for three different merchant good categories (MCGs), March 2021: (a) Food and groceries (b) restaurants (c) retail



(a)

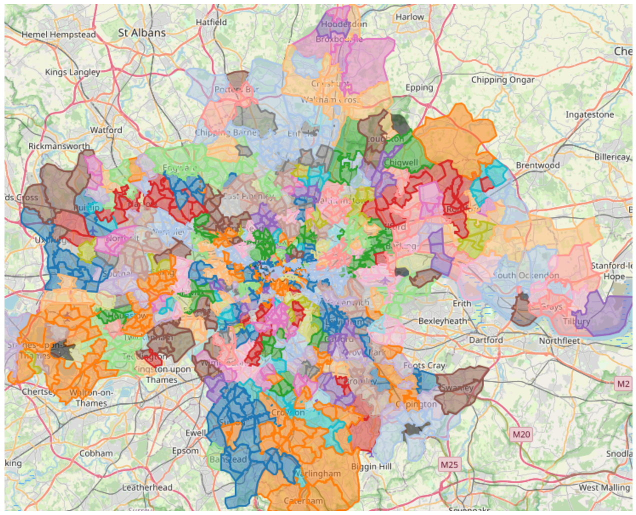


(b)

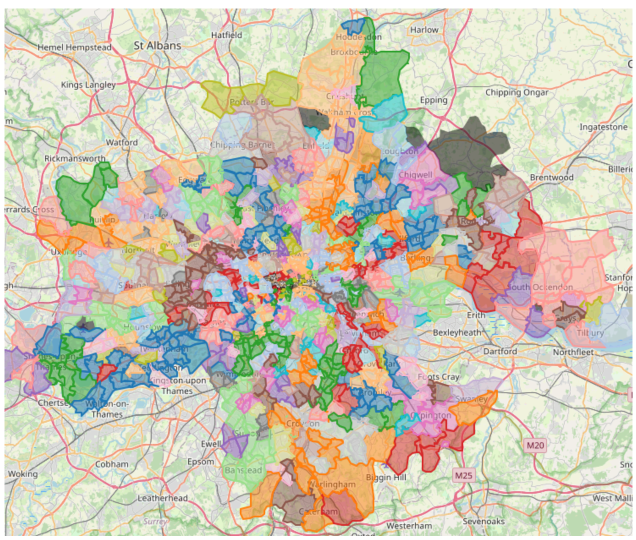


(c)

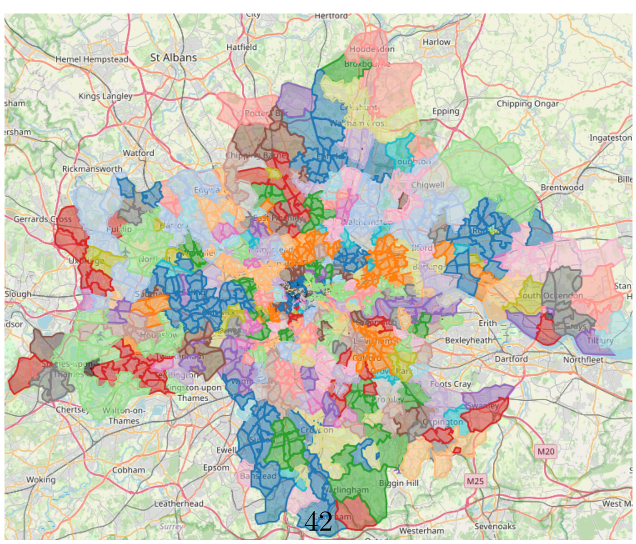
Figure B11: London cluster locations for three different time periods, food and retail: (a) March 2019 (b) March 2021 (c) March 2024



(a)



(b)



(c)

C Demographics and market size

We regress geographical market size on the demographic characteristics discussed above on the market area size, for each of our MCGs, in March 2024. This is a log-log regression specification, so the coefficients represent the percentage increase in the market size, once the proportion of each category within the demographic characteristic increases by one percent. We regress each demographic characteristic separately and then together in the final model. For illustrative purposes, we show here apparel, automotive, food and groceries and restaurants

Figure C1: Number of cardholder postal sectors, by MCG category, March 2024

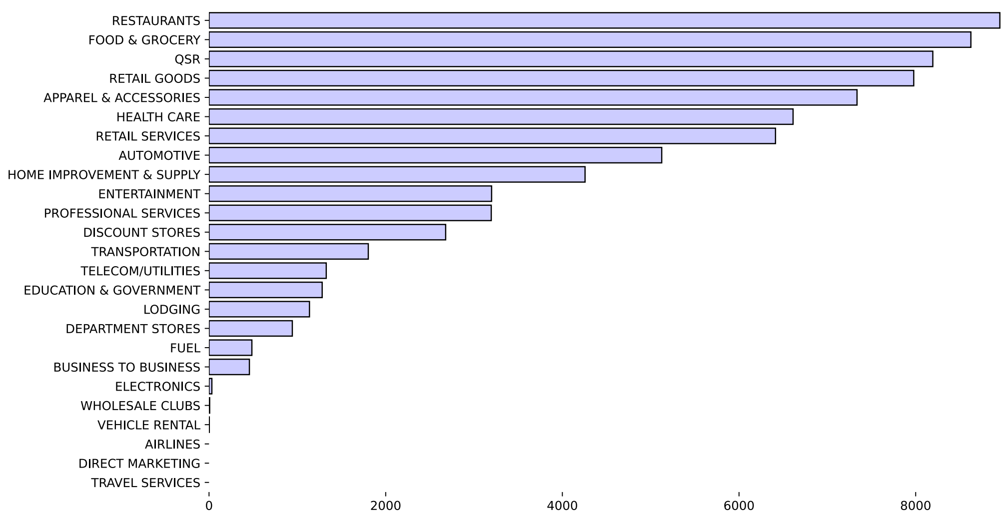
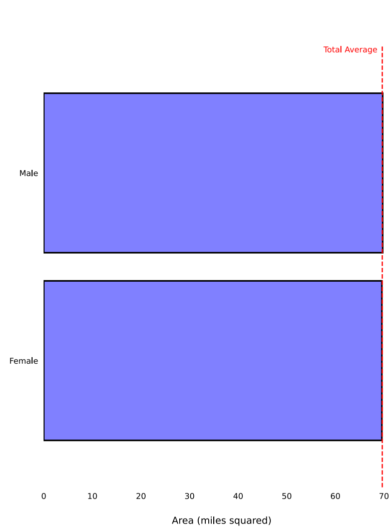
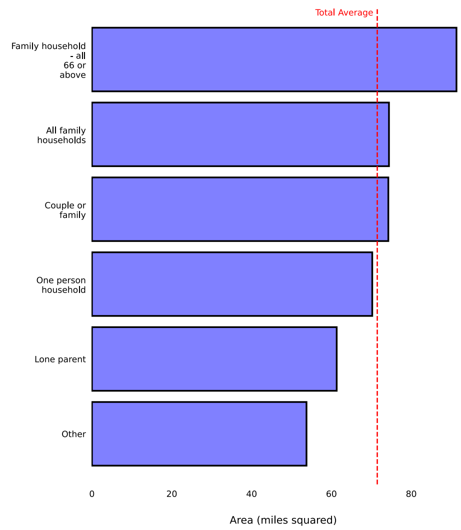


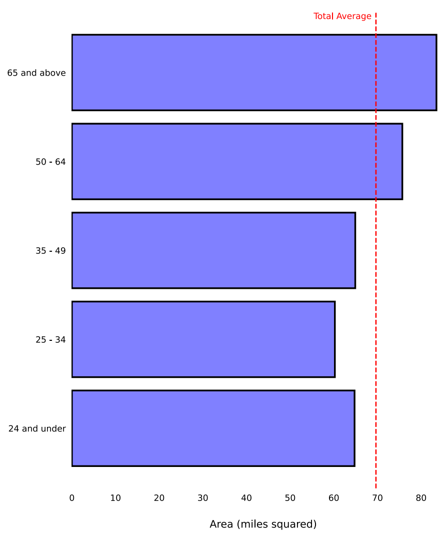
Figure C2: Average market size (in square miles miles), in food and groceries, based on demographic characteristics around: (a) gender (b) household composition (c) age, (d) accommodation type



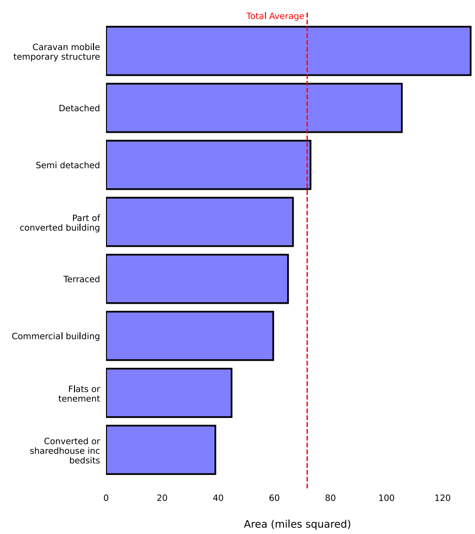
(a)



(b)



(c)



(d)

Regression of market size on demographic characteristics, log-log specification, March 2024, Apparel and Accessories

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	1.591***							-1.731
	(0.303)							(1.594)
Unemployed	-2.534***							-0.241
	(0.185)							(0.156)
Economically inactive	2.147***							-1.294*
	(0.215)							(0.745)
Students	X							X
24 and under		-0.280						0.122
		(0.389)						(0.649)
25 - 34		-1.293***						0.303
		(0.372)						(0.298)
35 - 49		-2.255***						-0.724
		(0.571)						(0.443)
50 - 64		5.750***						-0.463
		(0.781)						(0.408)
65 and above		-0.833*						1.834***
		(0.440)						(0.562)
Detached			0.886***					0.098
			(0.100)					(0.063)
Semi Detached			-0.205					-0.089
			(0.145)					(0.069)
Terraced			0.304***					0.098*
			(0.116)					(0.052)
Flats			-0.530***					-0.034
			(0.086)					(0.054)
Converted			-0.088					0.010
			(0.081)					(0.033)
Part of converted building			0.490***					0.002
			(0.084)					(0.031)
Commercial building			-0.096					-0.166***
			(0.113)					(0.042)
Mobile home			0.224***					0.034**
			(0.039)					(0.014)
Asian; Asian-British				-0.011				-0.034

				(0.097)				(0.036)
Black; Black-British; Carribean, African				-0.076				0.107***
				(0.110)				(0.035)
Mixed; multiple ethnic groups				-0.571***				-0.045
				(0.215)				(0.072)
White				1.972***				-0.303
				(0.113)				(0.187)
Other ethnicity				-0.238**				-0.035
				(0.119)				(0.036)
Work from home					0.503***			0.642***
					(0.172)			(0.211)
Metro					-0.010			0.016
					(0.047)			(0.015)
Train					-0.380***			-0.001
					(0.062)			(0.022)
Bus					-0.391***			-0.018
					(0.083)			(0.038)
Car					1.866***			0.206
					(0.362)			(0.346)
Other commuting modes					X			X
Population density						-0.725***		-0.798***
						(0.047)		(0.017)
Expenditure per card-holder							0.868**	-0.015
							(0.373)	(0.070)
Car ownership								X
Occupation								X
Gender								X
Family structure								X
Education								X
Observations	414	415	412	414	395	414	473	395
R^2	0.733	0.694	0.793	0.685	0.783	0.370	0.011	0.988

Note: *Stars denote statistical significance obtained from estimating clustered standard errors, at the market level, with stars indicating *** p<0.01, ** p<0.05, * p<0.10*

Regression of market size on demographic characteristics, log-log specification, March 2024, Automotive

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	1.335***							0.488*
	(0.160)							(0.278)
Unemployed	-2.226***							-0.064**
	(0.094)							(0.032)
Economically Inactive	2.159***							-0.159
	(0.131)							(0.155)
Students	X							X
24 and under		0.338*						-0.108
		(0.199)						(0.132)
25 - 34		-1.142***						-0.041
		(0.220)						(0.064)
35 - 49		-2.345***						-0.151
		(0.314)						(0.094)
50 - 64		4.337***						-0.086
		(0.401)						(0.093)
65 and above		-0.145						-0.110
		(0.231)						(0.122)
Detached			0.961***					0.001
			(0.047)					(0.013)
Semi Detached			-0.132*					0.011
			(0.067)					(0.016)
Terraced			0.234***					0.006
			(0.054)					(0.011)
Flats			-0.533***					-0.011
			(0.045)					(0.011)
Converted			-0.015					0.002
			(0.041)					(0.007)
Part of converted building			0.294***					-0.020***
			(0.040)					(0.006)
Commercial building			-0.021					-0.014
			(0.055)					(0.008)
Mobile home			0.185***					-0.001
			(0.020)					(0.003)
Asian; Asian-British				0.040				-0.024***

				(0.059)				(0.008)
Black; Black-British; Carribean, African				-0.217***				-0.002
				(0.060)				(0.008)
Mixed; multiple ethnic groups				-0.268**				0.016
				(0.125)				(0.017)
White				1.554***				-0.094***
				(0.070)				(0.035)
Other ethnicity				-0.121				0.001
				(0.077)				(0.009)
Work from home					0.529***			-0.066
					(0.090)			(0.045)
Metro					-0.078***			0.004
					(0.023)			(0.003)
Train					-0.272***			0.006
					(0.032)			(0.005)
Bus					-0.500***			-0.016**
					(0.041)			(0.008)
Car					0.940***			-0.147*
					(0.183)			(0.075)
Other commuting modes					X			X
Population density						-0.845***		-0.968***
						(0.019)		(0.005)
Expenditure per card-holder							1.113***	0.014
							(0.103)	(0.009)
Car ownership								X
Occupation								X
Gender								X
Family structure								X
Education								X
Observations	1187	1188	1166	1187	1139	1187	1335	1121
R^2	0.616	0.572	0.721	0.494	0.714	0.621	0.080	0.996

Note: *Stars denote statistical significance obtained from estimating clustered standard errors, at the market level, with stars indicating *** p<0.01, ** p<0.05, * p<0.10*

Regression of market size on demographic characteristics, log-log specification, March 2024, Food and Grocery

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	0.877***							-0.680*
	(0.105)							(0.365)
Unemployed	-2.193***							-0.041
	(0.069)							(0.050)
Economically Inactive	2.256***							-0.399**
	(0.081)							(0.193)
Students	X							X
24 and under		0.880***						0.559***
		(0.123)						(0.180)
25 - 34		-0.828***						0.343***
		(0.130)						(0.088)
35 - 49		-2.371***						-0.093
		(0.217)						(0.133)
50 - 64		3.101***						0.461***
		(0.283)						(0.122)
65 and above		0.344**						0.608***
		(0.161)						(0.141)
Detached			0.867***					0.027
			(0.033)					(0.018)
Semi Detached			-0.176***					-0.022
			(0.044)					(0.021)
Terraced			0.090**					0.039**
			(0.037)					(0.017)
Flats			-0.313***					-0.033*
			(0.030)					(0.019)
Converted			-0.019					-0.003
			(0.027)					(0.011)
Part of converted building			0.274***					-0.018*
			(0.027)					(0.010)
Commercial building			-0.019					-0.065***
			(0.038)					(0.013)
Mobile home			0.251***					0.010*
			(0.015)					(0.005)
Asian; Asian-British				0.125***				-0.025*

				(0.041)				(0.014)
Black; Black-British; Carribean, African				-0.130***				0.044***
				(0.043)				(0.014)
Mixed; multiple ethnic groups				-0.395***				-0.103***
				(0.092)				(0.029)
White				1.735***				0.065
				(0.049)				(0.054)
Other ethnicity				-0.197***				-0.007
				(0.054)				(0.016)
Work from home					0.592***			0.402***
					(0.065)			(0.079)
Metro					-0.075***			-0.005
					(0.017)			(0.006)
Train					-0.325***			-0.007
					(0.024)			(0.009)
Bus					-0.481***			-0.019
					(0.033)			(0.016)
Car					0.592***			0.296***
					(0.114)			(0.101)
Other commuting modes					X			X
					(0.107)			(0.039)
Population density						-0.909***		-0.819***
						(0.015)		(0.007)
Expenditure per card-holder							1.058***	0.022*
							(0.077)	(0.013)
Car ownership								X
Occupation								X
Gender								X
Family structure								X
Education								X
Observations	2236	2242	2151	2237	2115	2227	2518	2045
R^2	0.647	0.615	0.754	0.574	0.739	0.629	0.070	0.980

Note: *Stars denote statistical significance obtained from estimating clustered standard errors, at the market level, with stars indicating *** p<0.01, ** p<0.05, * p<0.10*

Regression of market size on demographic characteristics, log-log specification, March 2024, Restaurants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	1.191***							-3.385***
	(0.137)							(0.660)
Unemployed	-2.260***							-0.105*
	(0.086)							(0.058)
Economically Inactive	2.007***							0.084
	(0.100)							(0.272)
Students	X							X
24 and under		0.768***						0.028
		(0.209)						(0.236)
25 - 34		-1.138***						-0.167
		(0.162)						(0.110)
35 - 49		-3.143***						-0.037
		(0.281)						(0.162)
50 - 64		4.693***						-0.068
		(0.357)						(0.160)
65 and above		-0.250						0.665***
		(0.205)						(0.229)
Detached			0.900***					0.113***
			(0.040)					(0.027)
Semi Detached			-0.350***					-0.058*
			(0.054)					(0.030)
Terraced			0.217***					0.073***
			(0.044)					(0.023)
Flats			-0.518***					-0.020
			(0.033)					(0.020)
Converted			0.027					0.009
			(0.033)					(0.014)
Part of converted building			0.352***					0.010
			(0.031)					(0.012)
Commercial building			-0.032					-0.077***
			(0.043)					(0.016)
Mobile home			0.160***					0.004
			(0.016)					(0.006)
Asian; Asian-British				-0.061				0.004

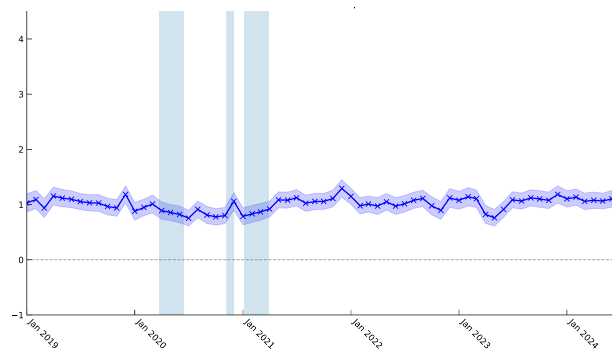
				(0.051)				(0.016)
Black; Black-British; Carribean, African				-0.085*				0.052***
				(0.051)				(0.015)
Mixed; multiple ethnic groups				-0.636***				-0.094***
				(0.111)				(0.034)
White				1.665***				0.103
				(0.060)				(0.083)
Other ethnicity				-0.079				0.014
				(0.058)				(0.017)
Work from home					0.469***			0.207**
					(0.074)			(0.095)
Metro					-0.095***			-0.003
					(0.020)			(0.007)
Train					-0.285***			-0.012
					(0.026)			(0.009)
Bus					-0.401***			0.006
					(0.035)			(0.015)
Car					0.967***			0.014
					(0.143)			(0.139)
Other commuting modes					X			X
Population density						-0.595***		-0.822***
						(0.016)		(0.009)
Expenditure per card-holder							2.671***	0.250***
							(0.177)	(0.047)
Car ownership								X
Occupation								X
Gender								X
Family structure								X
Education								X
Observations	1754	1762	1711	1756	1618	1747	1982	1590
R^2	0.527	0.535	0.647	0.454	0.629	0.439	0.103	0.970

Note: *Stars denote statistical significance obtained from estimating clustered standard errors, at the market level, with stars indicating *** p<0.01, ** p<0.05, * p<0.10*

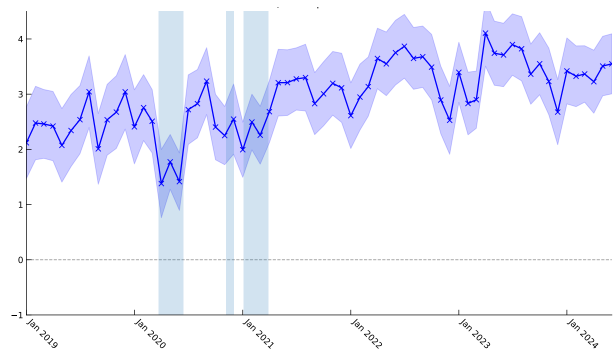
C.1 Coefficients across time

We perform this regression every for each month in our dataset. This allows us to assess whether the relationship between market size and demographics has changed, for instance during Covid-19 or the cost of living crisis. We see that expenditure per cardholder, our proxy for wealth, had less of an effect on market size faced by a consumer during Covid-19, for Food and Grocery shopping, Quick Service Restaurants, Retail Goods, and Retail Services.

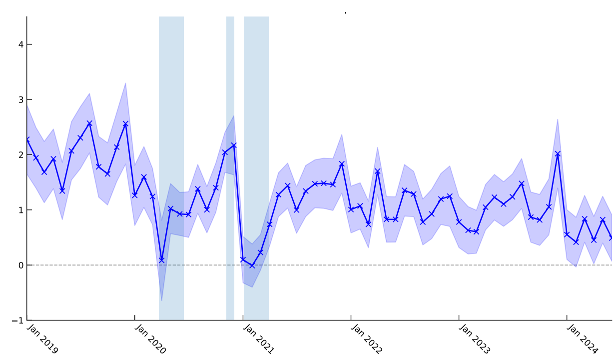
Figure C3: Coefficient of regressing market size against expenditure per cardholder, UK: (a) Food and Grocery (b) Quick Service Restaurants (QSR) (c) Retail Goods (d) Retail Services



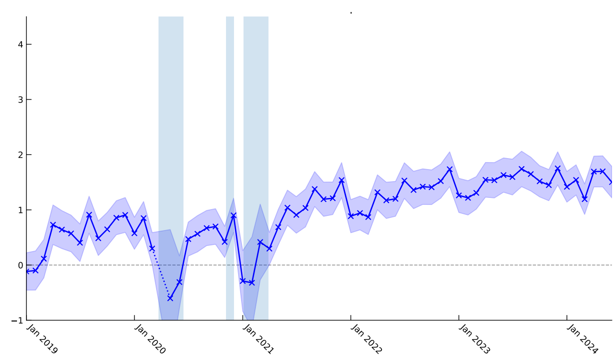
(a)



(b)



(c)



(d)

D Threshold selection

To select an appropriate threshold, we compare the market size obtained by our clustering algorithm to external estimates of how far people travel to make purchases. We therefore calculate our market's geographic size, and the spend within and across markets, for different threshold levels.

We analyse data from the Department for Transport's (DfT) National Travel Survey (DfT, 2023). The data is collected at an annual frequency. DfT collect data through face-to-face interviews and a seven-day self-completed travel diary. The survey is voluntary and covers around 16,000 individuals and 7,000 households in England, per year. The product sectors in this data do not align perfectly with the Merchant Category Groups (MCGs) in our data. As a result, we were only able to make comparisons for certain MCGs. For the remaining MCGs, we chose a default threshold of 0.4, corresponding to the mode of the thresholds for MCGs for which a comparison with DfT data was possible.

We calculate the distribution of how far people travel to buy goods and services, for each sector, on average across Covid (2020 and 2021) and non-Covid years (2019, 2022, and 2023) separately.

Our Visa data is in the tables below. We have selected the data that is most corresponds to the sectors in the National Transport Survey trip purpose categories. Note that the Within spend metric is the proportion (between 0 and 1) of spend occurring within the same cardholder market. The same applies for within transactions and within cardholders.

Based on these tables, and comparing with the Visa data, we select the following central thresholds for twelve MCGs in D25. For the remaining MCGs, which had negigible overlaps with the "trip purpose" categories in the NTS data, we select an arbitrary central threshold of 0.4, which is the most modal threshold value across MCGs for which we have external data. These central thresholds are used for clustering in all the above analysis, unless stated otherwise.

Figure D1: Distributions (in miles) of how far people travel from their home to go shopping, during Covid years, by their trip purpose category: (a) Food shopping (b) Non-food shopping (c) Public activities (d) Personal business - eat and drink (e) Drink with friends (f) Personal business - medical (g) Personal business - other (h) Other social activities (j) Other activities

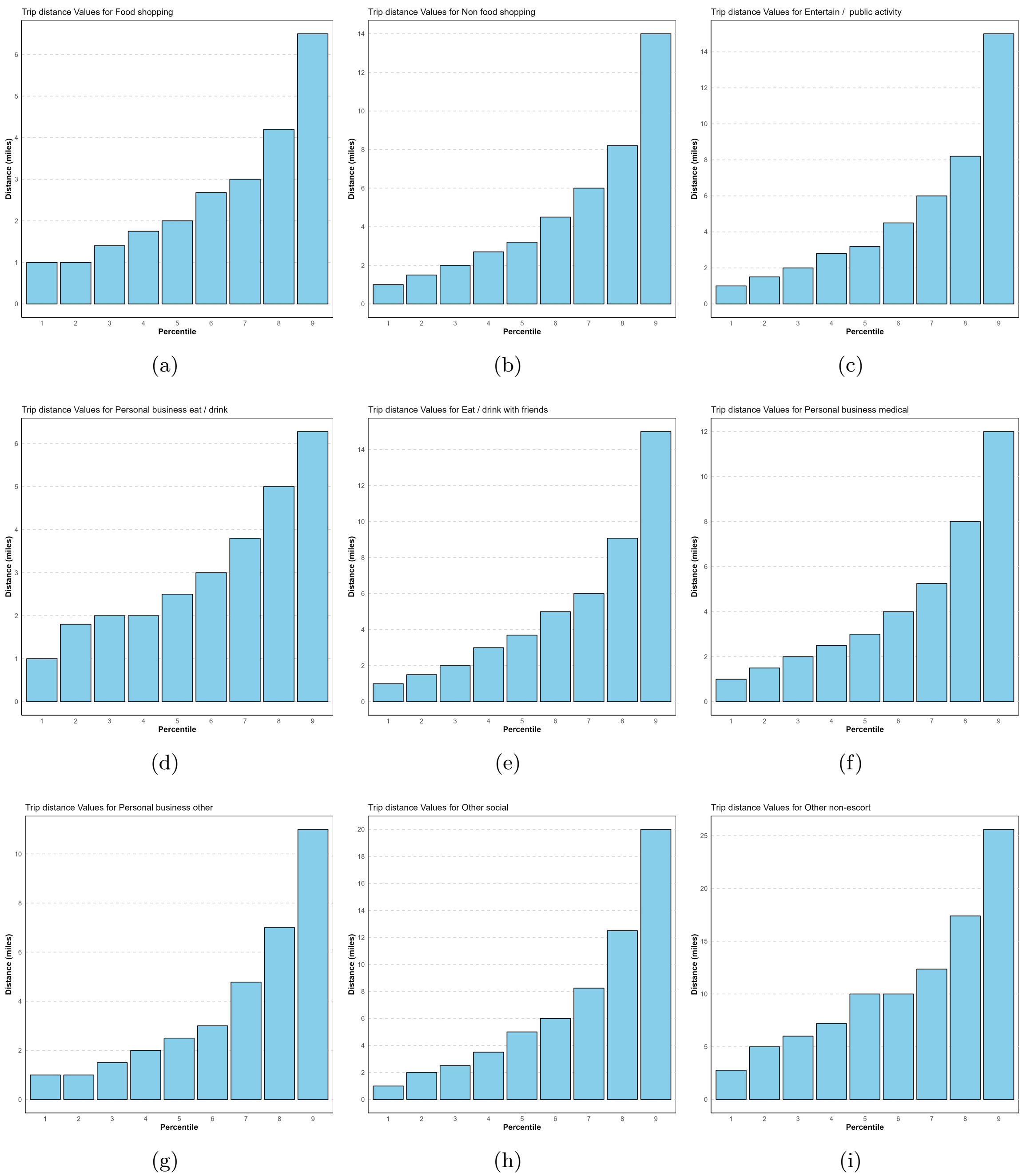


Figure D2: Distributions (in miles) of how far people travel from their home to go shopping, during non-Covid years, by their trip purpose category: (a) Food shopping (b) Non-food shopping (c) Public activities (d) Personal business - eat and drink (e) Drink with friends (f) Personal business - medical (g) Personal business - other (h) Other social activities (j) Other activities

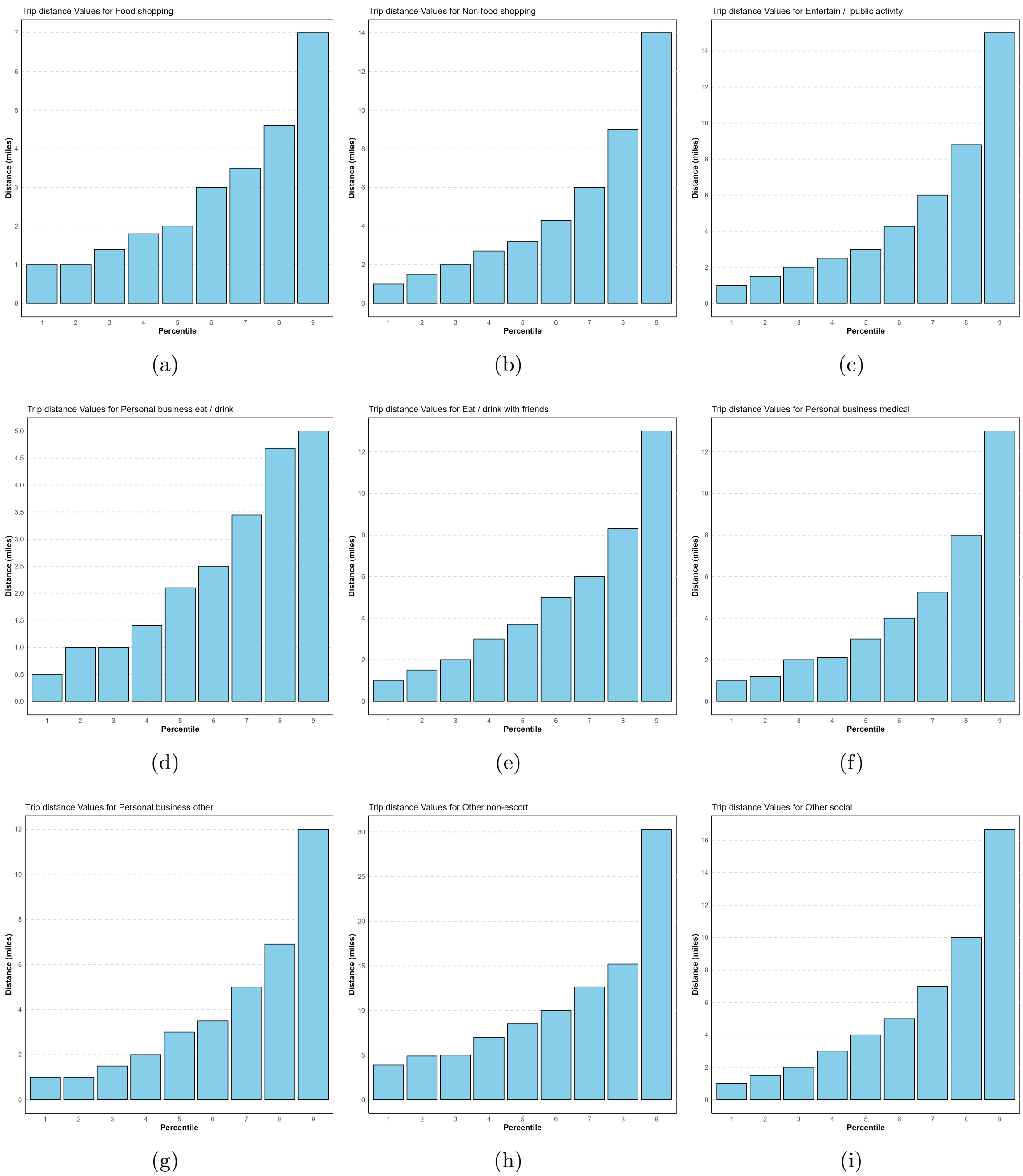


Table D1: Apparel and Accessories, Covid

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	17.0	17.3	17.1	16.4	16.2	15.9	16.0	16.4
Average area	74.2	136.6	206.4	330.4	464.3	748.9	1229.6	2712.0
Within spend	0.0	0.1	0.3	0.4	0.5	0.6	0.6	0.7
Within transactions	0.0	0.1	0.3	0.4	0.5	0.6	0.7	0.8
Within card	0.0	0.1	0.3	0.4	0.5	0.6	0.7	0.8
80 Percentile distance	33.1	34.3	31.9	25.7	23.4	22.8	20.5	19.8
80 Percentile area	245.2	449.7	641.2	810.7	1103.4	1528.5	2493.5	4201.0
Count	1660.0	962.0	613.0	412.0	289.0	192.0	114.0	50.0

Table D2: Apparel and Accessories, Non-Covid

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	16.5	16.7	16.6	15.7	15.1	15.1	15.1	15.9
Average area	78.0	134.5	217.2	325.4	462.2	697.7	1120.9	2660.9
Within spend	0.0	0.1	0.3	0.4	0.5	0.6	0.6	0.7
Within transactions	0.0	0.1	0.3	0.4	0.5	0.6	0.7	0.8
Within card	0.0	0.1	0.3	0.4	0.5	0.6	0.7	0.8
80 Percentile distance	32.5	32.9	30.0	24.7	22.4	21.8	20.0	20.0
80 Percentile area	249.4	434.3	681.7	762.6	1068.6	1450.4	2360.3	4255.8
Count	1611.0	944.0	607.0	418.0	290.0	194.0	119.0	53.0

Table D3: Automotive, Covid

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	3.7	3.9	4.0	4.0	4.1	4.2	4.2	4.2
Average area	20.2	34.4	54.7	77.5	108.0	139.4	218.9	389.1
Within spend	0.1	0.2	0.4	0.5	0.5	0.6	0.7	0.8
Within transactions	0.1	0.2	0.4	0.5	0.6	0.6	0.7	0.8
Within card	0.1	0.2	0.4	0.5	0.6	0.6	0.7	0.8
80 Percentile distance	7.1	7.2	7.2	7.1	7.2	7.3	7.2	6.6
80 Percentile area	70.1	122.4	167.8	238.1	292.4	352.6	479.8	772.5
Count	3205.0	2183.0	1552.0	1166.0	913.0	711.0	492.0	294.0

Table D4: Automotive, Non-Covid

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	3.7	3.9	4.0	4.0	4.1	4.2	4.2	4.2
Average area	20.2	34.4	54.7	77.5	108.0	139.4	218.9	389.1
Within spend	0.1	0.2	0.4	0.5	0.5	0.6	0.7	0.8
Within transactions	0.1	0.2	0.4	0.5	0.6	0.6	0.7	0.8
Within card	0.1	0.2	0.4	0.5	0.6	0.6	0.7	0.8
80 Percentile distance	7.1	7.2	7.2	7.1	7.2	7.3	7.2	6.6
80 Percentile area	70.1	122.4	167.8	238.1	292.4	352.6	479.8	772.5
Count	3205.0	2183.0	1552.0	1166.0	913.0	711.0	492.0	294.0

Table D5: [Department Stores, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	7.0	5.7	5.2	5.1	5.0	5.0	4.9	4.7
Average area	237.3	354.5	409.2	410.2	425.7	432.1	497.8	885.6
Within spend	0.0	0.0	0.2	0.6	0.6	0.7	0.8	0.8
Within transactions	0.0	0.0	0.3	0.5	0.7	0.7	0.8	0.9
Within card	0.0	0.0	0.3	0.5	0.7	0.7	0.8	0.8
80 Percentile distance	17.3	13.8	9.8	9.7	9.3	9.2	8.9	8.7
80 Percentile area	1026.5	1162.5	1206.9	1206.9	1206.9	1442.6	1639.2	1729.1
Count	580.0	389.0	326.0	311.0	299.0	288.0	253.0	183.0

Table D6: [Department Stores, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	6.7	5.8	5.2	5.1	5.0	5.0	4.8	4.5
Average area	268.4	374.6	382.8	394.4	416.3	416.3	464.0	705.3
Within spend	0.0	0.0	0.4	0.6	0.6	0.7	0.7	0.8
Within transactions	0.0	0.0	0.4	0.6	0.6	0.8	0.8	0.9
Within card	0.0	0.0	0.4	0.6	0.6	0.8	0.8	0.9
80 Percentile distance	17.8	13.2	10.3	9.9	9.7	9.7	9.5	8.1
80 Percentile area	1056.8	1196.6	1226.1	1468.8	1469.4	1469.4	1494.8	1619.3
Count	548.0	370.0	319.0	307.0	299.0	285.0	253.0	189.0

Table D7: [Discount Stores, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	9.0	6.9	5.8	5.3	5.2	5.1	4.9	4.8
Average area	105.2	148.2	180.5	188.1	213.9	259.0	350.8	781.7
Within spend	0.0	0.0	0.7	0.7	0.8	0.8	0.8	0.9
Within transactions	0.0	0.0	0.7	0.7	0.8	0.8	0.9	0.9
Within card	0.0	0.0	0.6	0.7	0.7	0.8	0.8	0.9
80 Percentile distance	21.5	18.7	12.1	8.8	8.6	8.1	8.1	7.4
80 Percentile area	379.0	500.2	509.9	513.3	597.9	709.0	883.2	1445.8
Count	1153.0	717.0	511.0	441.0	392.0	330.0	249.0	148.0

Table D8: [Discount Stores, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	8.6	6.5	5.4	5.0	5.0	4.8	4.7	4.8
Average area	108.8	146.0	165.5	173.6	203.0	241.8	364.2	745.0
Within spend	0.0	0.0	0.7	0.7	0.8	0.8	0.8	0.9
Within transactions	0.0	0.0	0.7	0.8	0.8	0.8	0.9	0.9
Within card	0.0	0.0	0.6	0.7	0.7	0.8	0.8	0.9
80 Percentile distance	21.5	18.7	11.8	8.7	8.4	8.1	7.6	7.6
80 Percentile area	356.0	515.6	568.3	568.3	598.9	675.0	950.4	1382.7
Count	1112.0	682.0	504.0	435.0	394.0	336.0	252.0	153.0

Table D9: [Electronics, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.3
Average area	33.0	33.0	33.0	33.0	33.0	33.0	33.0	33.0
Within spend	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Within transactions	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Within card	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
80 Percentile distance	2.3	2.3	2.3	2.3	2.3	2.3	2.4	2.4
80 Percentile area	64.4	64.4	64.4	64.4	64.4	64.4	64.4	64.4
Count	41.0	41.0	41.0	41.0	41.0	41.0	40.0	38.0

Table D10: [Electronics, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
Average area	42.9	42.9	42.9	42.9	42.9	42.9	42.9	42.9
Within spend	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Within transactions	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Within card	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
80 Percentile distance	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4
80 Percentile area	64.4	64.4	64.4	64.4	64.4	64.4	64.4	64.4
Count	40.0	40.0	40.0	39.0	39.0	39.0	39.0	37.0

Table D11: [Entertainment, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	20.6	20.0	19.3	18.3	17.8	17.4	17.8	17.6
Average area	44.2	82.0	138.0	196.3	276.1	347.1	573.5	1630.2
Within spend	0.0	0.0	0.2	0.4	0.5	0.6	0.6	0.7
Within transactions	0.0	0.0	0.2	0.4	0.5	0.6	0.7	0.8
Within card	0.0	0.0	0.2	0.4	0.5	0.5	0.6	0.8
80 Percentile distance	42.1	38.7	32.9	29.9	27.9	26.9	27.2	24.8
80 Percentile area	203.1	350.8	492.9	560.9	774.4	1065.1	1666.9	4009.4
Count	1857.0	1108.0	712.0	519.0	409.0	308.0	202.0	105.0

Table D12: [Entertainment, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	19.8	19.4	18.2	17.7	17.5	16.8	16.7	17.1
Average area	41.7	94.2	154.5	208.4	285.9	430.4	562.9	1518.3
Within spend	0.0	0.0	0.2	0.4	0.5	0.6	0.6	0.7
Within transactions	0.0	0.0	0.2	0.4	0.5	0.6	0.7	0.8
Within card	0.0	0.0	0.2	0.4	0.5	0.5	0.6	0.8
80 Percentile distance	41.1	37.6	31.4	27.3	26.2	25.9	25.2	24.2
80 Percentile area	213.0	363.7	511.7	655.7	841.7	1067.4	1643.2	3224.8
Count	1851.0	1109.0	718.0	528.0	419.0	317.0	214.0	112.0

Table D13: [Food and Grocery, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	5.7	5.5	5.4	5.2	5.1	5.1	5.0	4.9
Average area	10.4	16.7	23.4	31.7	39.9	51.4	85.2	211.0
Within spend	0.2	0.3	0.4	0.5	0.6	0.6	0.7	0.8
Within transactions	0.2	0.4	0.5	0.5	0.6	0.6	0.7	0.8
Within card	0.1	0.2	0.3	0.3	0.4	0.4	0.5	0.6
80 Percentile distance	13.7	12.8	11.2	10.0	9.5	9.2	8.7	8.4
80 Percentile area	58.3	83.2	107.8	134.7	161.5	216.7	297.6	534.1
Count	4899.0	3756.0	3024.0	2499.0	2045.0	1616.0	1134.0	595.0

Table D14: [Food and Grocery, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	5.8	5.7	5.4	5.3	5.3	5.2	5.2	5.0
Average area	10.4	16.2	24.8	33.2	42.5	54.7	83.3	222.7
Within spend	0.1	0.3	0.4	0.5	0.5	0.6	0.7	0.8
Within transactions	0.1	0.3	0.4	0.5	0.5	0.6	0.7	0.8
Within card	0.1	0.2	0.3	0.3	0.3	0.4	0.5	0.6
80 Percentile distance	14.2	13.2	11.6	10.8	10.3	9.9	9.6	8.3
80 Percentile area	60.1	83.7	109.5	134.1	167.3	221.0	322.9	555.9
Count	4793.0	3662.0	2903.0	2373.0	1951.0	1533.0	1089.0	570.0

Table D15: [Home Improvement and Supply, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	5.2	5.3	5.2	5.1	5.0	5.0	5.0	4.9
Average area	36.3	54.6	78.4	98.8	128.9	166.7	240.6	419.7
Within spend	0.0	0.1	0.3	0.4	0.5	0.6	0.7	0.8
Within transactions	0.0	0.1	0.2	0.4	0.5	0.6	0.7	0.8
Within card	0.0	0.1	0.2	0.4	0.5	0.6	0.7	0.8
80 Percentile distance	9.8	9.7	9.2	8.5	8.2	8.0	7.7	7.2
80 Percentile area	122.8	190.8	245.6	295.9	339.8	499.8	682.5	976.4
Count	2716.0	1786.0	1261.0	980.0	779.0	580.0	386.0	224.0

Table D16: [Home Improvement and Supply, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	5.1	5.1	5.0	4.9	4.8	4.8	4.6	4.5
Average area	35.4	54.3	79.2	100.4	124.0	153.9	252.3	444.9
Within spend	0.0	0.1	0.2	0.4	0.5	0.6	0.7	0.8
Within transactions	0.0	0.0	0.2	0.4	0.5	0.6	0.7	0.8
Within card	0.0	0.0	0.2	0.3	0.4	0.6	0.7	0.8
80 Percentile distance	9.5	9.4	8.7	8.3	7.9	7.6	7.5	6.7
80 Percentile area	128.9	202.3	263.8	297.0	337.3	487.6	673.4	971.7
Count	2722.0	1811.0	1259.0	992.0	804.0	604.0	399.0	226.0

Table D17: [QSR, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	10.3	10.9	11.2	11.2	11.3	11.4	11.7	11.7
Average area	16.7	29.0	48.1	71.0	95.0	166.1	292.0	710.8
Within spend	0.1	0.2	0.3	0.4	0.4	0.5	0.6	0.7
Within transactions	0.0	0.2	0.3	0.3	0.4	0.5	0.6	0.7
Within card	0.0	0.2	0.2	0.3	0.4	0.4	0.5	0.6
80 Percentile distance	19.8	20.8	20.3	19.1	18.2	18.1	17.3	16.4
80 Percentile area	59.2	97.4	141.2	206.1	331.2	463.2	666.1	1358.2
Count	3520.0	2351.0	1659.0	1254.0	937.0	646.0	398.0	176.0

Table D18: [QSR, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	10.2	10.8	10.8	11.1	11.3	11.3	11.5	11.5
Average area	16.8	30.9	49.8	72.4	98.3	182.8	339.3	736.9
Within spend	0.0	0.2	0.3	0.4	0.4	0.5	0.6	0.7
Within transactions	0.0	0.2	0.3	0.4	0.4	0.5	0.6	0.7
Within card	0.0	0.1	0.2	0.3	0.4	0.4	0.5	0.6
80 Percentile distance	19.8	20.6	20.0	18.8	18.6	17.9	17.6	16.4
80 Percentile area	60.2	110.9	150.2	219.9	313.1	459.6	659.6	1629.6
Count	3425.0	2296.0	1612.0	1220.0	909.0	634.0	391.0	175.0

Table D19: [Restaurants, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	15.0	16.0	16.8	17.3	18.2	18.8	19.2	19.9
Average area	14.1	23.8	36.1	54.5	80.8	151.8	284.3	723.9
Within spend	0.1	0.2	0.3	0.3	0.4	0.4	0.5	0.6
Within transactions	0.2	0.2	0.3	0.4	0.4	0.5	0.5	0.6
Within card	0.1	0.2	0.2	0.3	0.3	0.4	0.4	0.5
80 Percentile distance	23.8	24.8	25.2	24.9	25.2	25.6	25.6	24.6
80 Percentile area	43.0	66.8	97.3	150.6	226.6	369.4	552.2	1246.7
Count	3854.0	2696.0	1990.0	1468.0	1073.0	737.0	429.0	180.0

Table D20: [Restaurants, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	14.3	15.4	16.0	16.5	17.3	18.1	18.6	18.7
Average area	14.5	24.5	39.3	55.0	88.7	161.2	288.9	723.0
Within spend	0.1	0.2	0.3	0.3	0.4	0.4	0.5	0.6
Within transactions	0.2	0.3	0.3	0.4	0.4	0.5	0.6	0.7
Within card	0.1	0.2	0.2	0.3	0.3	0.4	0.4	0.5
80 Percentile distance	22.8	23.9	24.1	23.8	24.3	24.5	24.4	24.1
80 Percentile area	43.6	68.1	99.5	154.4	226.7	363.9	552.7	1152.5
Count	3775.0	2616.0	1932.0	1449.0	1060.0	718.0	423.0	177.0

Table D21: [Retail Goods, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	10.6	11.1	11.4	11.2	11.4	11.4	11.4	11.1
Average area	39.3	67.9	96.2	137.2	206.5	324.3	476.5	1021.0
Within spend	0.1	0.3	0.3	0.4	0.5	0.6	0.7	0.8
Within transactions	0.1	0.3	0.4	0.5	0.6	0.6	0.7	0.8
Within card	0.1	0.3	0.4	0.5	0.5	0.6	0.7	0.8
80 Percentile distance	21.2	21.5	20.3	18.6	17.5	16.4	14.5	14.5
80 Percentile area	106.5	171.0	256.7	338.8	522.2	702.4	974.6	1967.4
Count	2448.0	1572.0	1074.0	777.0	563.0	387.0	231.0	113.0

Table D22: [Retail Goods, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	10.3	10.8	11.1	11.0	11.0	10.8	11.1	10.8
Average area	42.2	67.6	94.1	146.3	184.9	331.5	497.2	954.3
Within spend	0.1	0.3	0.4	0.4	0.5	0.6	0.7	0.8
Within transactions	0.1	0.3	0.4	0.5	0.6	0.6	0.7	0.8
Within card	0.1	0.3	0.4	0.5	0.5	0.6	0.7	0.8
80 Percentile distance	20.7	20.8	19.4	18.1	16.9	15.4	15.0	13.8
80 Percentile area	112.7	189.6	264.0	350.8	476.6	703.8	1018.4	2030.1
Count	2402.0	1535.0	1049.0	759.0	554.0	376.0	231.0	114.0

Table D23: [Retail Services, Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	3.1	3.2	3.2	3.3	3.4	3.5	3.5	3.5
Average area	13.1	22.6	30.9	43.3	58.0	82.1	125.2	242.3
Within spend	0.2	0.4	0.5	0.5	0.6	0.7	0.8	0.8
Within transactions	0.2	0.4	0.5	0.6	0.6	0.7	0.8	0.9
Within card	0.2	0.4	0.5	0.6	0.6	0.7	0.8	0.8
80 Percentile distance	5.8	6.1	5.8	5.8	6.0	6.0	6.0	5.9
80 Percentile area	55.1	74.4	97.8	118.4	145.8	199.9	320.3	455.5
Count	3509.0	2426.0	1810.0	1415.0	1138.0	863.0	608.0	362.0

Table D24: [Retail Services, Non-Covid]

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Average distance	3.0	3.1	3.1	3.2	3.3	3.4	3.5	3.5
Average area	12.8	21.7	29.6	42.5	55.5	80.8	120.6	223.9
Within spend	0.2	0.4	0.5	0.5	0.6	0.7	0.8	0.8
Within transactions	0.2	0.4	0.5	0.6	0.6	0.7	0.8	0.8
Within card	0.2	0.4	0.5	0.6	0.6	0.7	0.8	0.8
80 Percentile distance	5.7	5.7	5.5	5.5	5.7	5.7	5.8	5.5
80 Percentile area	56.0	77.3	97.6	117.8	141.3	196.5	287.1	459.0
Count	3465.0	2395.0	1800.0	1423.0	1160.0	880.0	628.0	373.0

Table D25: Manually selected central thresholds used in our analysis.

	Covid	Non-Covid
Healthcare	0.4	0.4
Restaurants	0.4	0.4
Quick-service restaurants	0.4	0.4
Retail goods	0.3	0.3
Retail services	0.3	0.3
Discount stores	0.4	0.4
Entertainment	0.5	0.5
Home improvement and supply	0.4	0.4
Professional services	0.3	0.3
Food and grocery	0.4	0.4
Apparel and accessories	0.4	0.4
Automotive	0.4	0.4