Retail Centre Geographies:
Who, *What*, *Where* and *How*.

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy by:

Patrick Ballantyne

Department of Geography and Planning
University of Liverpool
United Kingdom

December 2022

Statement of Originality

I, **Patrick Ballantyne**, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, this has been indicated in the text.

Date: **09/12/2022**

Signature: **P. BALLANTYNE**

Abstract

Retail centres, the main cores of retailing in urban areas, have been of interest to geographers for decades, owing to their importance in the wider economic system, and role as the economic, social and cultural heart of local communities. However, in the 21st century, physical spaces of consumption such as retail centres face profound challenges, as the global retail sector continues to be threatened by wider economic pressures (e.g., recessions), the rising popularity of online shopping, changing consumer behaviour (e.g., experience retail) and recent ‘shocks’ such as the COVID-19 pandemic and cost-of-living crisis. As a result, there is significant evidence to suggest that physical sites of consumption both in the UK and U.S. are facing an ‘apocalypse’, as evidenced by increasing vacancies, reductions in footfall and the increasing share of sales online. Throughout this thesis, it is argued that by strengthening understandings of the geographies of retail centres, in particular *where* they are located, *what* characteristics they have and *who* uses them, it becomes easier to understand existing levels of provision and “vibrancy” in the retail (centre) system, and quantify their response to external pressures (e.g., COVID-19), facilitating better evidence-led decisions about retail location. Policy action is critical to protect the ‘brick-and-mortar’ component of the retail sector, but to be feasible there needs to be a comprehensive understanding of the geographies of retail centres at the national level, to identify how and where effective responses are needed, and assess their impact.

Thus far, whilst the spatial location (the *where*), characteristics (the *what*) and patronage (the *who*) of retail centre agglomerations have been of great interest, such insights have been relatively inconsistent, often focusing on specific locations (e.g., UK), or considering only a singular aspect of retail centre geographies. Significant progress has been made in delineating retail centre boundaries helping to answer *where* they are located, constructing classifications and hierarchies to understand *what* characteristics they have and developing catchments that capture *who* is using them and where they are coming from. However, these efforts have typically been concentrated in the UK or in rich local case studies, providing significant scope for national-level insights in other international locations. This is of great significance, given the power of retail centre geographies as a tool for understanding wider retail sector processes; a large body of work has shown *how* retail centre geographies can be used as geographic data tools to better understand the response to external phenomena; the 2008 recession, rising popularity of online shopping and COVID-19, to name a few. However, for such important

insights to be feasible, it is argued that comprehensive understandings of retail centre geographies at the national level are needed, particularly capturing *where* they are located, *what* characteristics they have and *who* uses them, all of which can be used as powerful tools to demonstrate *how* the wider retail sector is responding to these pressures.

Addressing these research gaps, this PhD thesis aims to progress understandings of the *who, what*, *where* and *how* of retail centre geographies, presented across three empirical chapters, all published in highly reputable journal articles. In particular, analytical and conceptual frameworks are developed which yield the first comprehensive understanding of the geographies of retail centres in the U.S., comprising retail centre boundaries, a two-tier retail centre typology and accompanying set of retail catchments, for the Chicago Metropolitan Statistical Area (chapter three) and national extent of the U.S (chapter four). In chapter five the utility of retail centre geographies is demonstrated, showing *how* they can be used to better understand how the retail sector is recovering from the COVID-19 pandemic, through examination of trajectories of retail centre recovery (and decline), using an unstable dataset derived from mobile phone applications. Whilst subject to limitations, the findings of this PhD thesis have contributed to the use of new data and methods, theoretical, conceptual, and substantive knowledge about retail centre geographies, and generated significant implications for the development of public policy and future research about the *who, what, where* and *how* of retail centre geographies.

Publications

This thesis is comprised of three empirical chapters, which have been published as regular research papers in highly reputable journals (see below). A number of additional projects have taken place during this PhD, resulting in additional research outcomes listed below.

Thesis Publications

- Ballantyne, P., Singleton, A., Dolega, L., Credit, K., 2021. A framework for delineating the scale, extent and characteristics of American retail centre agglomerations. *Environment and Planning B: Urban Analytics and City Science.* 49 (3) pp 1112-1128. DOI: https://doi.org/10.1177/23998083211040519.
- Ballantyne, P, Singleton, A., Dolega, L., Macdonald, J., 2022. Integrating the who, what and where of American retail centre geographies. *Annals of the American Association of Geographers,* DOI: https://doi.org/10.1080/24694452.2022.2098087.
- Ballantyne, P., Singleton, A., Dolega, L. Using unstable data from mobile phone applications to examine recent trajectories of retail centre recovery. *Urban Informatics* (accepted).

Other Publications

- Ballantyne, P.J., Singleton, A.D., Dolega, L. 2021. A regional exploration of retail visits during the COVID-19 pandemic. *Regional Studies, Regional Science.* 8 (1), pp 366-370. DOI: https://doi.org/10.1080/21681376.2021.1973548.
- Ballantyne, P. 2020. Review of *An introduction to R for spatial analysis and mapping (Second Edition)* by Chris Brunsdon and Lex Comber. *International Journal of Geographical Information Science.* 34 (1), pp 202-203. DOI: https://doi.org/10.1080/13658816.2019.1647541.
- Ballantyne, P., Sanderson, R., Davies, A., Houlden, V., Maclachlan, A., Owen, G., Robinson, C., Shin, H., Wartmann, F., Williams, I., Wilkin, J. "No formula for life and career": Early-career experiences of quantitative human geographers. *The Geographical Journal* (accepted).

Acknowledgements

Undertaking a PhD amidst a global pandemic has presented significant challenges; nearly two years of mobility and travel restrictions have altered what is a 'typical' PhD experience. I feel that without the support of some key people, I would not have been able to successfully complete this thesis, so at this point I would like to pay tribute to some of these amazing people, and thank them for their help and support.

Firstly, to Prof. Alex Singleton, for his continuous academic (and personal) support throughout my PhD. Over the past four years, Alex has been incredibly kind, and ensured that my skills are refined and targeted at specific projects, to ensure I could complete my studies in time, whilst listening to my own interests and giving me plenty of ownership over my project. Alex has also given me lots of encouragement and opportunities to expand my academic portfolio, network and career, going above and beyond as a PhD supervisor, dedicating significant time to issues I have had over the past four years. I really appreciate the time and effort he put into this PhD project, helping shape me into the researcher I am today, and for that I am truly grateful.

Secondly, to my secondary supervisor Dr. Les Dolega. Les has always been on hand to offer support and advice on life in general, as well as how best to frame my research projects within the field of retail geography, often turning a paper/chapter draft around with constructive comments in a matter of days. I am particularly grateful to Les for his encouragement during my time as a Research Assistant and in the early months of my new role at Liverpool, which has given me the confidence to continue my academic journey at Liverpool as a Postdoctoral Researcher, and my interests in retail geography.

Thirdly, to my other co-authors Jacob Macdonald and Kevin Credit, as well as to the early-career panellists, Rachael Sanderson, and members of the QMRG committee, whom I worked with on a written commentary for the Geographical Journal. These people were instrumental in helping to push further publications over the line with their expertise and advice, but also gave me the opportunity to collaborate externally with other academics outside of Liverpool, which has opened a number of opportunities since, and inspired me to continue my academic journey.

Thanks to all at the CDRC, SafeGraph and Geolytix who have provided data that has been invaluable to my PhD. In particular, a big thanks to SafeGraph, who through their 'Data for

Academics' programme made the primary objective of this PhD thesis feasible, following months of searching for useful and accessible sources of data on retailer locations for the U.S.

I also want to thank all of my colleagues at the Geographic Data Science Lab and CDRC, who have offered me general support, advice, and friendship, reviewed drafts of papers and been there to listen when I am facing significant challenges in my research. The GDSL has been a fantastic place to do a PhD, with each and every member having an impact on my overall PhD experience. There are too many people to mention here, but I particularly wanted to acknowledge the contributions of Alessia Calafiore, Martin Fleischmann, Maurizio Gibin (CDRC), Ollie O'Brien (CDRC), May Chen (CDRC), Natalie Rose, Olivia Horsefield, Dan Owen, Sian Teesdale, James Murphy, Krasen Samardzhiev, Nikos Patias, Michael Mahony, Maya Middleton-Welch, Francisco Rowe, Dani Arribas-Bel, Mark Green, Caitlin Robinson, Jeremiah Nieves, Ron Mahabir and Olga Gkountouna, all of whom have contributed in some way to ensuring my PhD is a success.

I also wanted to thank some of my closest friends, for their undying support over the past four years. Particular thanks to Tilly Clough, Ella Bytheway-Jackson, Libby Hopkins, John O'Neill, Jess Cerasale, Elaine Reilly and Clare Bohan, who have been there at various stages of my academic journey. Thank you to all of you for listening, discussing, supporting, and looking out for me over the past four years, and being there to celebrate my major achievements.

A final thanks goes to Mum and Dad. As the first of the family to go to university and first to complete a PhD, I want to thank you both for your undying support and understanding over the last 8 years, especially when I was facing significant personal issues or struggling with the pressures of doing a PhD in a pandemic. Your love and care during this time has given me the opportunity to flourish in my academic career without any major concerns or stresses, and I will be eternally grateful for this, especially given the challenges you already face with Liam. Unfortunately, Dad you can't retire just yet (haha), but I hope to one day return the favour to you both.

Authors Declaration

Below are a series of formal declarations from my two supervisors and other co-authors, giving their permission to use co-authored published material in this PhD thesis.

I, **Alex Singleton,** give permission for the following papers, co-authored with Patrick Ballantyne, to appear in the PhD thesis:

- Ballantyne, P., Singleton, A., Dolega, L., Credit, K., 2021. A framework for delineating the scale, extent and characteristics of American retail centre agglomerations. *Environment and Planning B: Urban Analytics and City Science.* 49 (3) pp 1112-1128. DOI: https://doi.org/10.1177/23998083211040519.
- Ballantyne, P, Singleton, A., Dolega, L., Macdonald, J., 2022. Integrating the who, what and where of American retail centre geographies. *Annals of the American Association of Geographers.* DOI: https://doi.org/10.1080/24694452.2022.2098087.
- Ballantyne, P., Singleton, A., Dolega, L. Using unstable data from mobile phone applications to examine recent trajectories of retail centre recovery. *Urban Informatics* (accepted).
- Ballantyne, P.J., Singleton, A.D., Dolega, L. 2021. A regional exploration of retail visits during the COVID-19 pandemic. *Regional Studies, Regional Science.* 8 (1), pp 366-370. DOI: https://doi.org/10.1080/21681376.2021.1973548.

Date: **09/12/2022** Signature:

I, **Les Dolega,** give permission for the following papers, co-authored with Patrick Ballantyne, to appear in the PhD thesis:

- Ballantyne, P., Singleton, A., Dolega, L., Credit, K., 2021. A framework for delineating the scale, extent and characteristics of American retail centre agglomerations. *Environment and Planning B: Urban Analytics and City Science.* 49 (3) pp 1112-1128. DOI: https://doi.org/10.1177/23998083211040519.
- Ballantyne, P, Singleton, A., Dolega, L., Macdonald, J., 2022. Integrating the who, what and where of American retail centre geographies. *Annals of the American Association of Geographers.* DOI: https://doi.org/10.1080/24694452.2022.2098087.
- Ballantyne, P., Singleton, A., Dolega, L. Using unstable data from mobile phone applications to examine recent trajectories of retail centre recovery. *Urban Informatics* (accepted).
- Ballantyne, P.J., Singleton, A.D., Dolega, L. 2021. A regional exploration of retail visits during the COVID-19 pandemic. *Regional Studies, Regional Science.* 8 (1), pp 366-370. DOI: https://doi.org/10.1080/21681376.2021.1973548.

Date: **09/12/2022** Signature: Les Dolega

I, **Kevin Credit**, give permission for the following paper, co-authored with Patrick Ballantyne, to appear in the PhD thesis:

- Ballantyne, P., Singleton, A., Dolega, L., Credit, K., 2021. A framework for delineating the scale, extent and characteristics of American retail centre agglomerations. *Environment and Planning B: Urban Analytics and City Science.* 49 (3) pp 1112-1128. DOI: https://doi.org/10.1177/23998083211040519.

Date: **09/12/22** Signature:

I, **Jacob Macdonald,** give permission for the following papers, co-authored with Patrick Ballantyne, to appear in the PhD thesis:

- Ballantyne, P, Singleton, A., Dolega, L., Macdonald, J., 2022. Integrating the who, what and where of American retail centre geographies. *Annals of the American Association of Geographers*. DOI: https://doi.org/10.1080/24694452.2022.2098087.

Date: **09/12/22** Signature: Jacob Macdonald

Table of Contents

List of Figures

Figure 23. Stability of devices between different UK regions, highlighting the increasing share of devices in retail centres located in London over the study period, and general stability of device decline in all other regions.

Figure 24. The functional response of retail centres visualised as the change in share of total devices (%) from the baseline, defined as the average share of devices (%) by retail centre type in August and September 2021.

Figure 25. The regional response of retail centres visualised as the change in share of total devices (%) from the baseline, defined as the average share of devices (%) by region in August and September 2021.

Figure 26. Recovery trajectories of Major Town Centres in the North-West (A) and District Centres in East of England (B). Trajectories have been calculated as the change in share of total devices (%) from the baseline, defined as the average share of devices (%) by region in August and September 2021.

Figure 27. Weekly visits to retail places in the Chicago MSA, from the week beginning 2nd March to the 6th April. Each iteration represents one week of data (e.g. WB 02/03), and the general trend seen across the iterations is a decrease in retail visits throughout the MSA in these early weeks, with the most dramatic decrease coinciding with the issuing of the ‘Stay at Home’ order (WB 16/03).

Figure 28. Weekly visits to retail places disaggregated by retail type. This figure highlights in particular the general increased uptake of ‘essential’ goods (e.g. convenience), following issuing of the ‘Stay at Home’ order (WB 16/03), and decreased usage of leisure-based retail (e.g. drinking places and restaurants). For information about the retail types and their aggregation see Table 13 and its accompanying caption.

List of Tables

Nomenclature

Below is a reference table that can be used to translate any abbreviations used in this thesis. Each abbreviation is defined in full in its first occurrence throughout the thesis, but following this only the abbreviation will be used (e.g., Central Business District, CBD hereafter).

Abbreviation	Definition
CBD	Central Business District
CDRC	Consumer Data Research Centre
CPT	Central place theory
DBSCAN	Density-based spatial clustering of applications with noise
DCLG	Department for Communities and Local Government
GIS	Geographic Information System
GPS	Global Positioning System
HDBSCAN	Hierarchical density-based spatial clustering of applications with noise
HSTF	High Streets Task Force
ICSC	International Council of Shopping Centres
IPM	Institute for Place Management
KDE	Kernel Density Estimation
LDC	Local Data Company
MSA	Metropolitan Statistical Area
OSM	OpenStreetMap
OSRM	Open Source Routing Machine
PAM	Partition around medoids
PCA	Principal Component Analysis
POI	Point of Interest

1. Thesis Introduction

1.1. Background

Geography, and quantitative geography in particular, remains highly distinctive and exceedingly vibrant (Castree et al., 2022), with younger cohorts continuing to emerge and produce exciting analytical research (Franklin, 2021), and as such there has arguably never been a better time to be a quantitative geographer or "geographic data scientist". Firstly, there has never been more abundant geographic data (Elwood et al., 2012), generated by many different sources including legacy ones like the decennial census and surveys, and new forms of data, such as those derived from mobile sensing platforms or earth observation technologies. This rapid increase in "Big Data" generates clear research opportunities, offering great potential for the (spatially enabled) social sciences to advance their understanding of a plethora of human and environmental issues (Singleton and Arribas-Bel, 2021). Secondly, the raw inputs needed to 'nourish' the field, such as open-source tools and computational capacity, are also in abundance (Franklin, 2021), enabling geographers to generate new insights and answer new questions. Thirdly, the methodologies used in the subdiscipline are now much more sophisticated (Johnston et al., 2019), whereas previously geographers were limited by the computational infrastructure available to them, restricting the implementation of complex algorithms or spatial analyses. As a result, quantitative geographers can now generate better understandings over space and time, by strengthening their links to data science and borrowing methodologies from the field (Singleton and Arribas-Bel, 2021), ensuring that geography remains important in an increasingly data driven, 'digital' world (Miller and Goodchild, 2015). Finally, from a philosophical and theoretical standpoint, quantitative geography has become very different (Johnston et al., 2019), with some supporting the concept of data-driven epistemology or the "Fourth Paradigm" (Kitchin, 2014), where data and computational insight replace theories, whilst others continue to adopt more nuanced and integrated approaches, establishing new conceptual and theoretical frameworks (Singleton and Arribas-Bel, 2021). As a result, quantitative geography is a great place to be for research in the 21st century as the world we live in continues to become more complex and dynamic, such that practitioners with shared interests have constructed lively research communities in which to produce high quality research, using new forms of data and advanced analytical techniques, to contribute new knowledge about some of the most important issues of our time (Castree et al., 2022).

The field of retail geography is one such research community; the geographies of retailing have been of great interest to geographers for a long time (Brown, 1992), coming to occupy a key role within the social-scientific agenda (Crewe, 2000), as evidenced by the increasing number of outlets in which to publish high quality (retail) research, and the presence of dedicated sessions at academic conferences. The formal field of retail geography began to grow in the 1930s and 1940s when the distinct connections between retail activities and urban morphology began to be recognised, followed in the mid-20th century with a series of pioneering research articles in retail geography by stalwarts such as Hoyt, Proudfoot and Huff in the U.S., and Christaller and Smailes in Europe. Towards the end of the 20th century, a more theoretically informed interrogation of retail activities emerged, in what was coined the "new retail geographies" by Wrigley and Lowe (1996), which described a reconstructed and theoretically engaged subdiscipline, that took its economic and cultural dimensions seriously, and argued that the transformation of retail capital and consumption spaces offered some of the most fascinating and challenging areas of study in human geography. Contemporary retail geography continues to engage with these historical debates, taking advantage of a greater availability of consumer data and open-source tools, improved computational capacity and conceptual and methodological innovations, to answer new questions about consumer behaviour and the contemporary retail environment.

However, answering such questions remains a significant challenge, as today's retail environment is inherently complex. The contemporary spaces of consumption we interact with have evolved through time, partially in response to changes in the planning of retail land-uses (Guy, 2007), but also in response to pressures and new technologies which have facilitated the rise of 'E-commerce', which continues to alter consumer behaviour, and subsequently the demands on consumption spaces and retailers (Singleton et al., 2016). These spaces of consumption, such as large shopping centres and high streets, fall within the definition of retail centres – "the main retail cores within urban areas" (Dolega and Celińska-Janowicz, 2015, p9); one of the most ubiquitous features of the commercial (and urban) environment, acting as centres for retail, but also associated activities such as leisure (Macdonald et al., 2022). The geographies of retail centres have been of interest to geographers for decades, particularly the spatial organisation of these geographical phenomena (i.e. *where* are they located), which can be traced back to the mid-20th century (Woodbury, 1928; Murphy and Vance, 1954). Retail geographers have also sought to understand the role of retail centres in the supply of consumers with goods and services, as well as their relative position in the wider retail (centre) system, through the creation of 'hierarchies' and 'typologies' that aim to present an understanding of *what* characteristics they have or *what* their role is within the wider retail (centre) system (e.g., International Council of Shopping Centres, ICSC hereafter, 2017; Brown, 1992; Dolega et al., 2021). Furthermore, retail geographers have also been interested in *who* interacts with these spaces of consumption and where they come from, through demarcation of retail (centre) catchments to better comprehend their role (and interactions) with demand and supply (e.g., Dolega et al., 2016; Lloyd and Cheshire 2017).

Owing to these advancements in the understandings of retail centre geographies, there has also been a significant research effort seeking to understand the response of retail centres to external pressures, such as economic shocks and crises (e.g., 2008 recession), 'E-commerce' and COVID-19, through examination of the changing economic performance of retail centres, and processes of retail decline (Enoch et al., 2022; Dolega and Lord, 2020). Such studies are vital, as the global retail sector continues to be challenged by economic shocks and crises, 'E-commerce', changing consumer behaviours (e.g., experiential retailing) and the recent COVID-19 pandemic and cost-of-living crisis, the latter of which has seen a rapid increase in interest rates, potentially destabilising lots of retail businesses (Grimsey, 2013). However, for such studies to be viable, it is important that a concrete understanding of the geographies of retail centres, specifically *where* they are located, *what* characteristics they have and *who* uses them is available. When obtained, taking advantage of new data and methods, these retail centre geographies can be used as geographic data tools to show *how* the retail sector is responding to these pressures.

1.2. Research Gaps

Existing research on retail centre geographies has made significant progress in advancing understandings about their existence and the ways in which they respond to changing consumer behaviour. However, despite significant research about retail centre geographies, in particular *where* they are located, *what* characteristics they have, *who* uses them and *how* they can be used as geographic data tools, there are still some key research gaps relating to international or 'global' understandings of retail centre geographies (*where*, *what* and *who*), and the use of new forms of data to quantify the impacts of new and emerging external pressures, which continue to present significant challenges for the economic performance of retail centres (e.g., COVID-19). These research gaps are described in more detail below.

The where of retail centre geographies.

Studies seeking to identify the location and boundaries of retail centres (the *where*) have sought to delineate their boundaries through use of high-resolution retailer location data and methodologies that can be used to simplify such data into polygons that represent their 'formal' boundaries. However, research efforts have typically suffered from a lack of geographic representation, or from being based on methodologies that are non-reproducible, built to suit the dataset or study area in question. For example, there has been an abundance of relevant research in the UK (e.g., Thurstain-Goodwin and Unwin, 2000; Pavlis et al., 2018; Macdonald et al., 2022), and non-academic examples (e.g., OS, 2019; Geolytix, 2020), but these have typically been concentrated in the UK setting, with limited efforts elsewhere. Furthermore, the delineation approaches are often based on bespoke methodologies, making them much more difficult to generate such understandings in other international settings (Pavlis et al., 2018), where methodologies are often constructed in a way to suit the location data (and its limitations), as well as the study area in question, limiting further (global) applications.

Furthermore, given the dynamic nature of retail centres which expand and contract over time (Pavlis et al., 2018), there is significant scope for reproducible approaches to boundary delineation that can be updated over time to reflect such changes, especially in new international settings, such as the U.S., utilising available data and developing innovative (yet reproducible) methodologies, to generate robust understandings about the spatial location of retail centres – the *where*. The need for such understandings is based on the idea that retail challenges in the UK are not unique, as the U.S. has faced significant challenges and undergone dramatic transformations in the way retailing is carried out; declining popularity of downtown department stores, increasing size of units, evolution of superstores, category killers and 'big-box' merchandisers, and most recently facing a swath of retail closures or 'retail apocalypse' (Helm et al., 2020; Wrigley and Lowe, 2002; Basker, 2007). Thus, given the dynamic nature of U.S. retailing, there is a clear need to understand its existing spatial provision.

The what of retail centre geographies.

In trying to unpack the characteristics of retail centres, and how these relate to other centres in the system, studies have historically constructed hierarchies to capture functional similarities and differences (e.g., ICSC, 2017), based on assumptions about supply and demand drawn from Walter Christaller's (1933) Central Place Theory (CPT hereafter). However, recent examples have argued for such understandings to be multidimensional, accounting for a wider range of factors to capture the functional roles of retail centres. Dolega et al. (2021) proposed a new analytical framework through which to non-hierarchically classify retail centres in the UK, which generated significant potential to comprehend the characteristics and functional roles of retail centres in new international settings. For example, the data-driven nature of their framework provides significant scope for the expansion of retail centre classifications and typologies into other international settings for the first time, provided suitable data is available to capture key domains that are used to classify retail centres (Dolega et al., 2021). Furthermore, the framework has the added advantage of not being reliant on expert knowledge to construct it, which restricts the extension of retail centre classifications and typologies into new international locations (e.g., U.S.), where expertise is not widely available. As a result, there is significant scope to utilise this framework to generate a comprehensive understanding of the functional differences and characteristics of retail centres – the *what* - in other international settings.

The who of retail centre geographies.

Retail catchments describe the "areal extent from which the main patrons of a store or retail centre will typically be found" (Dolega et al., 2016, p78), and are particularly useful at obtaining information about *who* is using retail centres and where they are coming from. Techniques for estimating retail catchments can be either deterministic or probabilistic, and attempt to account for both supply and demand (Birkin et al., 2010), in order to accurately model consumer patronage behaviours, and estimate the areal extent from which patrons to a retail centre are found. Whilst probabilistic modelling is generally favoured over deterministic techniques, they are limited in that they often model how consumer patronage ought to occur, rather than accounting for actual patronage behaviours. However, new approaches are emerging utilising emerging high-resolution consumer datasets (Newing et al., 2015), and advancements in their application for model calibration (Wang et al., 2016). Recent examples have used these non-conventional forms of data such as that obtained from loyalty cards or social media platforms to calibrate models that accurately capture real consumer patronage behaviours (e.g., Waddington et al., 2018; Davies et al., 2019; Lloyd and Cheshire, 2017). However, research utilising non-conventional forms of data such as that obtained from mobile phones, and advancements in catchment modelling techniques, to derive catchments for retail centres as opposed to stores, are sparse in the literature (Pratt et al., 2014; Dolega et al., 2016), especially in other international settings.

The how of retail centre geographies.

Retail centres are dynamic geographic entities (Dolega and Celińska-Janowicz, 2015), contracting and expanding over time in response to external pressures such as economic recessions, 'E-commerce', changing consumer preferences, out-of-town retail formats and the COVID-19 pandemic. These pressures have resulted in significant transformations to the types of goods and services available (i.e., supply), and the decisions made by consumers about where to shop (i.e., demand), as well as an overall decline of physical 'brick-and-mortar' consumption spaces, as evidenced by rising vacancy rates (Tselios et al., 2018), reduced footfall (High Streets Task Force, HSTF hereafter, 2021) and the increasing popularity of online shopping (Singleton et al., 2016). However, we do not fully understand the nature of these processes. There has been some research into the response of retail centres to external pressures, with studies exploring retail success and decline (Dolega and Lord, 2020; Jones et al., 2022), the resilience of retail centres to 'E-commerce' (Singleton et al., 2016) or the consequences of national lockdowns and public health restrictions on activity during the COVID-19 pandemic (Enoch et al., 2022; Frago, 2021; Trasberg and Cheshire, 2021). With the latter, research has typically focused on specific study areas (e.g., London) or the immediate response to the onset of the pandemic. Thus, there is a notable research gap in unpacking the longer-term impacts of COVID-19 on retail centres, particularly at the national extent, rather than through rich local case studies, seeking to understand spatio-temporal trends of recovery and/or decline.

1.3. Thesis Aims

Considering these emerging research gaps in retail centre geographies, three key research aims have been identified for this PhD thesis, to generate new insights about the *where*, *what*, *who* and *how* of retail centre geographies. An overview of each of these is presented below:

Aim One: Investigate whether recent advances in retail centre delineation and classification can be used to capture the geographies of retail centres into other international settings.

Chapter three seeks to utilise advancements in the delineation of retail centre boundaries (Pavlis et al., 2018) and the classification of retail centres (Dolega et al., 2021), to explore the potential for expansion of retail centre geographies, specifically the *where* and the *what,* into other international settings. Using data from SafeGraph, an overview of *where* retail centres are located and *what* characteristics they have is provided, specifically focusing on the Chicago Metropolitan Statistical Area (MSA hereafter) as a case study. In particular, an approach to retail centre delineation utilising Hierarchical-DBSCAN (HDBSCAN hereafter) provides a reasonable mechanism and alternative to Pavlis et al. (2018), through which to delineate retail centre boundaries, but is limited in expansion beyond the city or metropolitan scale. Furthermore, adoption and enhancement of the framework established in Dolega et al. (2021) generates a useful representation of the salient characteristics and functions of retail centres in Chicago Metropolitan Statistical, but requires further enhancement to capture local and national retail niches, and generate an overview of the American retail centre system.

The content of this chapter can be found within the following published journal article:

Ballantyne, P., Singleton, A., Dolega, L. and Credit, K. 2022. A framework for delineating the scale, extent and characteristics of American retail centre agglomerations. *Environment and Planning B: Urban Analytics and City Science, 49* (3): 1112-1128. Available at: https://doi.org/10.1177/23998083211040519.

Aim Two: Generate a comprehensive understanding of the geographies of a national retail centre system outside of the UK.

Chapter four provides a comprehensive overview of retail centre geographies for the U.S., building on the achievements of chapter three (Ballantyne et al., 2022a). Again, using data from SafeGraph, retail centre boundaries are delineated (the *where*) using a method based around the hexagonal spatial index H3, and a much more granular retail centre classification is obtained (the *what*), by customising and enhancing the retail centre classification framework (Dolega et al., 2021), to better capture niches in U.S. retail. Furthermore, catchments are estimated for these retail centres, by calibrating a traditional Huff model with mobility data from SafeGraph, to robustly estimate consumer patronage to retail centres (the *who*). Through development of a conceptual framework through which to better understand the geographies of a national retail centre, and accompanying empirical insights about the *who, what* and *where* of U.S. retail centres, and the connections between them, a comprehensive overview of the U.S. retail centre system is provided for the first time, which can be used to generate insights about the response of the system to external shocks.

The content of this chapter can be found within the following published journal article:

Ballantyne, P., Singleton, A., Dolega, L. and Macdonald, J. 2022. Integrating the who, what and where of US retail center geographies. *Annals of the American Association of Geographers.* Available at: https://doi.org/10.1080/24694452.2022.2098087.

Aim Three: Explore spatio-temporal trends of retail centre recovery using data derived from mobile phone applications.

Chapter five focuses on the use of retail centre geographies as geographic data tools, through which to better understand the longer-term response of retail centres to the COVID-19 pandemic (the *how*). Using new retail centre definitions for the UK and a large unstable mobility dataset from Geolytix, changes to activity within retail centres over a twelve-month period, characterised by the Omicron variant of COVID-19, are explored. In particular, significant focus is placed on how retail centres with different *functional* roles, regional geographies and *structural* characteristics have experienced significantly different recovery trajectories, through exploratory analysis and modelling of individual recovery trajectories. Furthermore, significant insights are contributed about the utility of the Geolytix mobility data for spatio-temporal analysis, highlighting its value when considered as snapshots and comparing trends between areas, rather than exploring temporal trends which are subject to changes in the devices uses to create the data. This research provides an initial basis upon which to use alternative forms of data (e.g., mobility data) to monitor ongoing processes of recovery and decline, providing evidence that can inform policy decisions and provide solutions to both acute and longer-term issues in the wider retail sector.

The content of this chapter can be found within the following published journal article, which has only recently been accepted:

Ballantyne, P., Singleton, A., and Dolega, L. Using unstable data derived from mobile phone applications to examine recent trajectories of retail centre recovery. *Urban Informatics* (accepted).

1.4. Overview

The remainder of this thesis is structured as follows. Chapter two provides an overarching review of the relevant literature, comprising key theoretical debates in retail geography and the main approaches that have been used in the past to understand the *where, what* and *who* of retail centres, before considering *how* they can be used as geographic data tools, and importantly *why* all of this matters. Chapters three, four and five comprise the empirical contributions of this thesis that accompany each of the three research aims. Finally, in chapter six, the achievements of this PhD thesis are reflected upon, discussing some of its notable findings and implications, considering the study limitations and suggesting how some of these insights will inspire future academic research about retail centre geographies.

2. Literature Review

2.1. Retail centres – the *place* of retailing

Retailing: a brief history

In modern history, the vast majority of retailing has occurred in stores or "outlets," defined as "any building where retailing is carried out, or where retail goods can be sold to the public from the premises without appointment" (Guy, 1998, p1), Historically however, retailing has occurred through more primitive means; early retail in 9000 BC existed in the form of the exchange of cows and sheep as goods, with early forms of currency and "Agoras" (marketplaces) appearing in Ancient Greece around 800 BC (Braun, 2015). Over time this trend of buying and selling goods developed, and physical stores began to emerge, with a transition from traditional stores like general stores, "mom-and-pops" and independents to department stores in the late 19th century (Braun, 2015). However, todays retail environment is much more complex than this, emerging as a result of transitions through a sequence of different retail formats (McArthur et al., 2016), with the sequence being different for each country, but largely similar for western countries such as the UK and U.S.

In the UK, the 20th century saw a number of new retail formats emerging, which were a product of changes to retail land-use planning and regulation. In particular, the last 30-40 years has seen the significant growth of out-of-town shopping centres and retail parks (Astbury and Thurstain-Goodwin, 2014), which provide consumers with greater convenience, plenty of parking, and 'one-stop' shopping, typically located away from traditional town centres (Dolega and Celińska-Janowicz, 2015). Shopping centres typically comprised large numbers of comparison goods (e.g. fashion retailers) with very limited provision of convenience goods (e.g. groceries), whilst the contemporary retail park has grown to offer a hybrid of these, between superstores (e.g., Home Bargains) and warehouses (e.g., TK Maxx). Further loosening of regulation at the same time also spawned a more market-led system, a period which also saw the emergence of discount stores (Burt et al., 2010). The result during this time was a significant decline of the traditional sites of retailing (i.e. high streets), which began to experience an erosion of market share (Wrigley et al., 2015), as consumers were provided with other options to fulfil their needs as consumers.

Across the Atlantic, the United States saw similar trends, where as a result of similar developments in market regulation, innovations in transport and an increasing suburban

American population, the retail system saw significant suburbanisation of retail capital and decline in downtown retailing (Wrigley and Lowe, 2002). This process of retail decline manifested itself in the declining popularity of downtown department stores, where the focus of retail was swept out to newly built regional shopping centres (Wrigley and Lowe, 2002). At the same time, the U.S. saw an increase in the size of retail units, with supermarkets becoming larger and evolving into 'superstores', whilst at the same time, large retail formats such as mass merchandisers (e.g., Wal-Mart) and category killers (e.g., Staples) began to emerge, providing a greater breadth and depth of products respectively (Wrigley and Lowe, 2002). These trends resulted in significant market consolidation, with "mom-and-pop" stores vanishing, and the market share of the sector being controlled by many of these "big-box" mass merchandisers and category killers (Basker, 2007). In recent years, American retail has been going through a "retail apocalypse", owing to large numbers of retail closures since 2015 (Helm et al., 2020), with experts predicting further contraction in the number of 'brick-and-mortar' retail outlets, stores and malls, with significant impacts on local communities (Helm et al., 2020).

The agglomeration of retailing

The spatial organisation of retailing has been of great interest to geographers for a long time, and agglomeration remains one of the key concepts in understanding the spatiality of retailing. Agglomeration refers to a mass or collection of things, and as a concept, its relevance to industrial structures has been studied widely. Human society and its industries have always seen spatially concentrated economic activities, which are a result of the benefits businesses can gain from location close to each other (Ellison et al., 2010), and the subsequent increases in economic productivity (Kerr and Kominers, 2015). For example, a clustering of heterogenous retail activities facilitates a multi-purpose shopping experience, which is both more appealing, and enables consumers to economise their overall shopping costs, through reductions in travel times (Reimers and Clulow, 2009). Furthermore, multi-purpose shopping increases the likelihood that a consumer will visit other stores in that agglomeration or retail centre, even if that was not their original intention (Lüer-Villagra et al., 2022); comparison shopping becomes a form a leisure. Finally, retailers benefit from spatial agglomeration as it builds productive competition; they each benefit from a stream of consumers visiting that area, competing for the share of the consumers time and money (Teller and Reutterer, 2008).

When defining an agglomeration of retail units then, these can be called “retail agglomerations” or “retail clusters” (Berman and Evans, 2013). However, in recent literature the term “retail centre” has been used to describe this phenomenon (Lloyd and Cheshire, 2017, Pavlis et al., 2018). Dolega and Celińska-Janowicz (2015, p9) provided a formal definition of retail centres as ‘the main cores of retailing in urban areas’. Given such (locational) emphasis on urban areas, retail centres also share many characteristics with other urban-economic phenomena, such as town centres, which are much more well defined and feature frequently in related retail (and urban) geography literature. For example, they can both be viewed as complex systems that constantly evolve (Thurstain-Goodwin and Unwin, 2000), expanding and contracting over time in relation to their relative attractiveness, market potential and competition (Pavlis et al., 2018). Furthermore, they are composed of a broad assortment of shops and associated activities, such as entertainment and leisure, which enrich the experience for consumers and visitors alike, improving the overall economic performance of these centres (Teller and Reutterer, 2008). However, it is important to highlight the key differences between town and retail centres, as they are arguably distinct urban-economic entities, and there is often confusion between the two. Town centres form the core of many urban areas in the UK, characterised by a clustering of socio-economic activities (Pavlis et al., 2018), comprising a retailing centre, a concentration of leisure/entertainment facilities, some services, a business sector and good transport accessibility (Coca-Stefaniak, 2013), and as such are the economic, social and cultural heart of the town (Haklay et al., 2001). Thus, we can often think of retail centres as existing within town centres, forming what is often a key part of the town centre’s offering and associated place product (ODPM and CASA, 2002).

However, there are many exceptions, such as those retail centres that occur away from town centres (Coca-Stefaniak, 2013), which includes retail parks, out-of-town shopping centres and strip malls in the U.S., which typically have very different characteristics. In particular, they have a predominant specialism in retail and are typically ‘purpose-built’, instead of growing and evolving over time (Guy, 1998), contrary to the ‘traditional’ UK town centre. However, it is not fair to exclude such purpose-built developments from definitions of retail centres, as despite different planning circumstances, their organisation as agglomerations of retail units still enables them to be recognised as distinct retail centres (Berman and Evans, 2013), thus this thesis would move to extend the definition initially proposed to be “the main cores of retailing and ancillary activities in geographical space”, thereby not excluding centres located in less urbanised areas, and accounting for their wider functional role.

Theoretical considerations

Thus, it is apparent that retail activities rarely exist in isolation, and this tendency of retail activities to agglomerate into clusters is considered an example of the centrality of retailing, an important concept in urban and economic geography (Latham et al., 2008). There are a variety of theoretical models that have tried to understand why the process of centrality occurs in urban areas, such as Burgess's concentric zone model and Hoyt's sector theory (Nong et al., 2019). However, one theory that has, and continues to be very useful in helping to understand the spatial structure of retail and service business is central place theory (CPT hereafter), which is arguably 'the widely accepted key model of retail organisation' (O'Brien and Harris, 1991).

CPT was introduced by Walter Christaller in 1933, originally being used to explain how settlements operated to provide goods to surrounding areas, based on special economic-geographic laws that determine the arrangement of settlements (Christaller, 1933). According to CPT, a 'central place' or urban centre will exist that serves the needs for goods and services in surrounding areas (Nong et al., 2019), through offering 'central place goods' that are commodities or services for which demand is dispersed across an area (Parr, 2017). Each central place (and its market area) will be centrally located with respect to the dispersion of demand, with consumers electing to be supplied by the closest location (Parr, 2017). Thus, CPT provides a conceptual model through which we can understand the retail location process at the centre level (Kohsaka, 1989). For example, it is common that the dominant or largest site supplying goods and services will be located at the cores of urban areas. Furthermore, the concept of a retail catchment, which is returned to later, is based on the premise that demand is dispersed across an area, as in CPT. Thus, Christaller's work on CPT provides a theoretical underpinning for helping to contextualise the supply of goods and services within urban areas, and some of the factors that determine their demand.

However, retail centres are not homogenous, occupying different functional roles, spanning different spatial scales and subject to different planning circumstances (Dolega et al., 2021), which arguably do not fit into Christaller's conventional hierarchies. However, the importance of CPT must not be overlooked, as it was instrumental in making empirical connections between function and scale, arguing that central places exist within a distinctive hierarchical structure (Parr, 2017); given that demand is determined by the frequency with which goods are purchased (Brown, 1992), a central place hierarchy could be constructed. CPT also established a clear difference between lower-order convenience and higher-order comparison goods (Brown, 1992), and Christaller also demonstrated how a central place could be ranked in terms of the population within an associated market area (Dennis et al., 2002). However, as a theoretical tool for understanding and classifying retail centres in today's contemporary retail environment, CPT has faced a significant amount of criticism. For example, many argue that CPT has become increasingly unrealistic, since it assumes that shoppers will always patronise their nearest centre (Brown, 1992). However, today movement is much easier, with significant reductions in the role of transport costs in determining demand (Parr, 2017). Furthermore, CPT assumes a uniform population distribution and statistic distribution of goods and services, which is problematic and unrepresentative in the polycentric cities of Great Britain, as argued by Dolega et al. (2016). A further complication involves the emergence of non-conventional retail channels such as 'E-commerce', which has significantly altered consumer behaviours (Dolega et al., 2021) and arguably does not fit into the systems outlined in CPT, as consumers can obtain their goods from anywhere in the world. Furthermore, the contemporary retail system is now much more complex; retailers may have mixed functional purposes or are 'omnichannel' in nature; where a hybrid of online and 'brick-and-mortar' retailing remains popular (Dolega et al., 2021). Thus, it can be said that the traditional hierarchy of urban systems, posited by CPT, is of limited contemporary utility because the relationship between retail centres in the wider retail system is far more complex.

This discussion of retail centres as 'central places' shares some significant overlaps with the work of Von Thünen (1826). In their theory of 'Agricultural Land Use', Von Thunen posited that a city will be located centrally within an 'isolated state' whereby farmers in the surrounding area are organised in such a manner as to maximise profits (O'Kelly and Bryan, 1996). The most intensive farming types like dairying and fruit and vegetable harvesting would be located in the nearest 'ring' to the city, and the least intensive, self-transporting farming types like animal products would be furthest away. Although this model is built on many unrealistic assumptions, it draws attention to the importance of the locational attributes of the property, which Von Thunen argued can be quantified as 'location rent' (O'Kelly and Bryan, 1996). This relates directly to some of the key ideas of Christaller's CPT, particularly the importance of distance to consumer patronage. Furthermore, some of these ideas remain largely applicable in modern retail environments where when considering consumer patronage behaviours, there is still a duality between the land rent – i.e., attributes of a location or retail centre – and the proximity to the market.

2.2. The *where* – the location, scale and extent of retail centres

The where: legacy delineations and insights

Understanding the location of retail centres, in particular their scale and extent requires formal definition or 'delineation' of their geographical boundaries, and has been of great interest since the early 20th century. The earliest related study was in 1928, where Woodbury seeked to identify the size of retail business districts in Chicago (Woodbury, 1928), using data supplied by physical maps and/or field observations, as was common in the 20th century. For example, Murphy and Vance (1954) identified the Central Business District (CBD hereafter) of U.S. cities based on the spatial information about retail and office premises, building heights, land values and pedestrian activity, and Clark (1967) carried out field observations in Christchurch, New Zealand, to identify the spatial extent of retail business centres based on a series of arbitrary thresholds, e.g., "associations of stores within at least 200 feet of each other".

However, in the late 1990s and early 2000s, work began on how new data and automated processes might be able to rigorously and systematically define the spatial organisation and morphology of retail centres (Thurstain-Goodwin and Unwin, 2000). In their paper, Thurstain-Goodwin and Unwin (2000) used kernel density estimations (KDE hereafter) to create continuous surface representations of four key characteristics of town centres; economy, property, diversity of use and visitor attractions. The outcome of this was an 'Intensity of Town Centredness surface, where analysis of peaks on the surface provided a means of delineating the spatial extent of town centres. The resulting town centre boundaries were then used by the Department for Communities and Local Government (DCLG hereafter) as the official town centre boundaries for the UK, proving the usefulness of such work.

The where: contemporary approaches to retail centre delineation

However, in recent years the delineation of such geographical phenomena has proven challenging as the factors used to determine the boundary often depend on the perspectives of the stakeholders (ODPM and CASA, 2002). A further challenge is that the morphology of town and retail centres continues to change and evolve over time (ODPM and CASA, 2002), making the delineation of their spatial extents especially challenging, calling for approaches that can provide systematic measures that can be updated over time. An additional constraint is that many highly accurate definitions of the spatial organisation of retail centres are created and held by commercial organisations such as Geolytix (Geolytix, 2020) and Ordnance Survey (OS, 2019), and as a result are not available to be used in research to answer geographical questions. However, the coupling of new forms of data and increasing availability of open data, alongside advancements in analytical capabilities and new conceptual and analytical frameworks through which to better understand such phenomena, have revived interest in the delineation of retail centre space (Dolega et al., 2016). In recent examples, there has been a preference for more explicit definitions using the locational attributes of individual stores (Porta et al., 2009; Han et al., 2019), rather than identification of the characteristic factors of retail (and town) centres, as in Thurstain-Goodwin and Unwin (2000). In particular, studies have started to emerge that use KDEs of location data to explore such agglomerations, as in China (Wang et al., 2014; Han et al., 2019), Italy (Porta et al., 2009), and the UK (Lloyd and Cheshire, 2017).

Recently however, there has been an emphasis on the applications of spatial cluster analysis to delineate the boundaries of retail centres. Spatial clustering algorithms such as density-based spatial clustering of applications with noise (DBSCAN hereafter) are used to assign classes, groups or 'clusters' to spatial data, having advantages over other algorithms like K-means in that they do not need a predetermined number of clusters, can identify arbitrarily shaped clusters and enable clustering based only on spatial similarity/dissimilarity between objects (McInnes and Healy, 2017; Campello et al., 2015). In their study, Pavlis et al. (2018) developed a modified version of DBSCAN, solving issues with heterogenous local point densities (Campello et al., 2013), to systematically delineate the location, scale and extent of retail centre boundaries in the U.K. Such definitions, using retailer location data from the Local Data Company (LDC hereafter), were made openly available as geographic data resources through the Consumer Data Research Centre (CDRC hereafter), being utilised in several further academic studies (Dolega et al., 2021; Comber et al., 2020; Trasberg and Cheshire, 2021). Furthermore, the authors made their R code available via a public GitHub repository, arguing that this provided an empirical basis upon which to delineate the location, scale and extent of retail centre boundaries in other international settings (Pavlis et al., 2018).

More recently, an updated version of the CDRC retail centre boundaries data product was developed (Macdonald et al., 2022), which utilised the H3 spatial indexing system and open-source data on retailer locations to derive 6,423 retail centres across the UK. The paper demonstrated the efficacy of H3 as a tool for understanding the spatial location of retail centres in the UK, but in particular made significant contributions by creating a methodology that is both robust and effective at scale, but also accessible and transparent enough to enable easy replication in other international settings. However, aside from these efforts by Macdonald et al. and Pavlis et al. (2018), limited efforts have been made to delineate retail centre boundaries in other international settings outside of the UK.

2.3. The *what* – positioning retail centres within the wider system

The what: historical classifications and hierarchies

Once the location, scale and extent of retail centres has been established, we can establish the different roles that they have, in particular considering how they relate to each other in the wider retail system. In order to generate such understandings, it is important to consider *what* characteristics they have, and how that relates to the characteristics of other retail centres in the system, traditionally through the development of classifications, hierarchies and typologies (Guy, 1998). These seek to understand the different retail agglomeration forms in the system and the roles that they occupy, and have a long legacy in the retail geography literature, but also remain prominent as retail continues to evolve and transform (Micu, 2019; Rao, 2020).

Historically, the system of (retail) centres was considered as a hierarchy, positing that retail is hierarchically organised (Brown, 1992), and can be classified based on assumptions about demand and supply from CPT. Such arguments were based on the idea that centres of differing scale (i.e., size) exhibited different functions to each other (Sadahiro, 2000), and such hierarchies have retained saliency in the present day, for example in the International Council of Shopping Centre's (ICSC hereafter) shopping centre classification (ICSC, 2017), which distinguishes American shopping centres based on their floor space and market areas. Whilst their classification does account for function (e.g., specialist centres, factory outlets), it is limited to purpose-built developments, excluding naturally evolved retail centres, such as downtown districts in the U.S., and high streets in the UK.

However, as discussed earlier, Christaller's traditional urban hierarchy of centres is no longer applicable when trying to conceptualise a system of retail centres. Recent literature has instead argued that we need to view the contemporary retail environment as a product of external and internal factors; it is multidimensional (Dolega et al., 2021). Furthermore, there is evidence to suggest that classifications that are non-hierarchical in nature, can better capture the functional differences between centres within them. One such example was proposed in 1992 by Brown, who demonstrated that retail centres could be distinguished non-hierarchically based on their form and function. Brown (1992) demonstrated that retail centres can exhibit a variety of forms, including clustered, linear or isolated ones, and that their functions can vary based on the types of retailing they offer; general, specialist or ‘ancillary’ (Brown, 1992).

The what: contemporary classifications and typologies

As discussed, the past few decades, has seen fundamental shifts in the way that people shop, reflecting the growing role of online channels, changing demands of consumers and need for experiential retail and leisure (Joseph and Kuby, 2016; Dolega and Lord, 2020). Thus, for any measure of retail function (i.e., classification) to have contemporary relevance, it has to reflect such changes, whilst also accounting for a greater ‘multidimensionality’ of descriptive input measures to effectively differentiate between different functions (Guy, 1998; Rao, 2020). Such approaches are becoming increasingly feasible, owing to advancements in the way classification is conceptualised and measured in retail geography (e.g. Brown, 1992), but also to advancements in analytical capacity and data (Dolega et al., 2021), providing significant scope for advanced classifications based on sophisticated empirical analysis (DeLisle, 2005).

In light of these advancements, a number of more sophisticated retail centre (and high street) classifications have begun to emerge. For example, Coca-Stefaniak (2013) demonstrated how a socioeconomic classification matrix could be used to unpack functional differences between high streets, based on the range of choices town centres have to make in order to attract visitors and the balance between social and economic profits needed to ensure a prosperous local high street. Another example was seen in Mumford et al. (2017), where footfall data from Springboard UK Limited and unsupervised machine learning techniques were used to verify the existence of four distinct monthly footfall signatures between UK high streets, and similar signatures at the daily scale. Furthermore, in Jones, Newing and Orford (2022), the authors captured a series of catchment characteristics (e.g., demographics) and town centre characteristics for Welsh town centres, and used these to construct a typology of towns in Wales, ranging from ‘Large Leisure Towns and Cities’ to ‘Small Independent Towns’.

Specifically for retail centres, and building on the location, scale and extent insights gained from Pavlis et al. (2018), Dolega et al. (2021) proposed that we needed to go further than just considering the form and function of centres. In their paper, they argued that data can be used to capture the multidimensionality of retailing, based on four key domains: composition, diversity, size and function and overall economic health. They gathered a series of variables spanning the four domains, and constructed a multidimensional typology, that described functional differences between UK retail centres, but did not organise these into a hierarchy. The authors argued that this multidimensional typology of retailing and service activity helps to understand how consumption spaces have transformed in recent years (Dolega et al., 2021), and the added benefit was that it can be re-constructed using available data, over numerous time periods and in different geographic locations, owing to its data-driven nature. However, setting aside the achievements of Dolega et al. in the UK, systematic nationwide and rigorous 'data-intensive' studies on the characteristics of retail centres are yet to be realised, especially in other international settings.

2.4. The *who* – retail centre catchments and the geography of consumer patronage

Retail (centre) catchments: an introduction

As geographers, we are always concerned with *who* is interacting with urban areas, and the same applies to retail centres. A term that is synonymous with such efforts is 'catchment', defined as an area that draws in a group of people, whether these be workers, customers or others (Lloyd & Cheshire, 2017). In retail specifically, the term retail catchment can be defined as the "areal extent from which the main patrons of a store or retail centre will typically be found" (Dolega et al., 2016, p1). Catchment analysis is an important tool for understanding, visualising and quantifying the extent of a market area, enabling insights into the spatial distribution or geography of consumer patronage (Segal, 1999), and in the case of retail centres, this involves asking *who* uses them and where they come from. Understanding differences in [retail] catchment areas is a useful exercise as it enables insights into how modern consumer behaviours vary, and what the main drivers behind demand might be (Waddington et al., 2018). Secondly, as argued by Halsall (2001), it is an important tool for decision making, as it helps us to better understand the effects of competition, and how it interacts with demand and supply. Thus, it is no surprise that understanding catchment areas is a key step for individual retailers when trying to determine the location of a new store (Dolega et al., 2016; Murad, 2005), since retailers will not locate new stores in areas where the market is already oversaturated. Catchments are however inherently complex, as they attempt to summarise the patronage of lots of different individual consumers, and the factors which affect these behaviours. In the simplest form, Birkin et al. (2010) posited that catchment extents are determined by two elements; supply and demand. With these in mind, there are a number of ways to approximate a catchment, with the use case in each circumstance dependent on the requirements of the study, availability of data and the analytical capability of the researcher (Dolega et al., 2016). However, broadly speaking, approaches can be split into two; deterministic and probabilistic.

Deterministic techniques

Deterministic catchment techniques are the simplest to implement; "those in which the values for the dependent variables of the system are completely determined by the parameters of the model" (Rey, 2015, p3). This implies that known input values are specified to generate the catchment area, such as the drawing of circular buffers around a store (or retail centre) based on distance, dependent on how far consumers are willing to travel (Segal, 1999; Murad, 2005; Berry et al., 2016). This method is described as being 'nominal' in the sense that the catchments are derived using a fixed distance from a defined point (Halsall, 2001). Although the rings can be fixed distance or variable distance between different stores or retail centres, what remains consistent is that the catchment area is circular (Segal, 1999). Similar in nature to concentric rings, 'drive-time' methodologies apply a nominal measure to derive catchment areas, in this case the time (or distance) it takes to drive from a store (or retail centre). The key difference from fixed-ring buffers is that this method uses digitised transport networks, speed limits and transport modes to generate a polygon that represents the extent to which a vehicle can travel from the store in all directions, under a certain time limit (Thompson & Walker, 2005; Rudavsky et al., 2009). A final deterministic method is 'anisotropic buffering', which uses less conventional shapes (e.g., thiessen polygons) to represent the directional and distance-related sensitivity of retail catchments. The outcome is similar to drive-time catchments in that each buffer will have a different distance in each direction (Mu, 2008; Widaningrum, 2015), arguably representing the catchments more realistically.

Perhaps the most obvious benefit of deterministic methods is that they are easy to conceptualise and implement (Dramowicz, 2005), often meaning they are cost-effective (Halsall, 2001), and as such, are a particularly useful tool for small independent retailers. A significant benefit of drive-time methods in particular is that they account for some of the logistical barriers facing consumers (e.g., traffic), which is something overlooked by many methods (Segal, 1999). The final benefit of these tools is that they can be combined with the outputs of spatial interaction models; a tool that will be introduced later, to plan for retail store locations more accurately

(Mu, 2008). More critically, there are however a significant number of drawbacks with these methods. Firstly, they often fail to account for overlapping catchment boundaries and competition (Dolega et al., 2016; Halsall, 2001), thus often over-estimating the market share of individual retailers (or retail centres). Secondly, consumers typically patronise more than one store, which is something not accounted for in these models either (Dolega et al., 2016), instead representing the store as having a monopoly over its catchment areas (Dramowicz, 2005). The final drawback is that they are rarely based on observed information or consumer trends, and as a result can be misleading (Halsall, 2001). Thus, it could be argued that these techniques over-simplify the complexity of consumer patronage behaviours, leading to catchment boundaries that are conceptually less robust than those derived from more sophisticated, probabilistic methodologies.

Probabilistic techniques

Probabilistic catchment techniques are characteristically different from deterministic ones in that they use information derived from observations and empirical data to generate catchments, treating the outcomes of the model, in this case the catchment area, as probability distributions rather than unique values (Rey, 2015). The first method in this probabilistic 'toolbox' is a gravity model, which apply Newtonian laws of physics to the modelling of shopper behaviour, and approximate a retail catchment by considering the spatial distribution of competing locations and evaluating their attractiveness to different groups (Huff, 1964; Segal, 1999). The theory was first applied to retail by Reilly (1931), who hypothesised that consumers trade off the cost of travel and the attractiveness of competing destinations in deciding where to shop, providing a series of breakpoints or distance decay curve, that represent the spatial interaction of these factors (Segal, 1999; Dolega et al., 2016). These early gravity models underpinned a much of the work on consumer patronage behaviours (Dramowicz, 2005), and triggered significant research interest into the prediction of consumer patronage. However, the early gravity models did have some significant drawbacks; they were difficult to develop for multiple stores or centres at one time (Dramowicz, 2005), requiring significant computational power, and typically considered attractiveness as a product of population and distance (Yriogen and Otero, 1998), which arguably overlooked many other factors that make a store (or retail centre) attractive.

As such, David Huff introduced the “Huff model” in 1964, as a new way to model consumer patronage behaviours. It calculated the probability that a consumer (i) would shop at a destination (j), with the focus being on the person not the store, since it is the people who determine the trade area of a store (Huff, 1964). Each probability was calculated on the theoretical basis that consumers are faced with a series of choices when deciding where to shop, opting for the greatest utility. Thus, Huff calibrated his model on three variables; distance, attractiveness and competition (Dramowicz, 2005); the distance parameter remained mostly unmodified, with some examples accounting for travel time/distance instead of Euclidean distance (Newing et al., 2015). Attractiveness was typically calculated using some combination of a series of factors, including store size, number of retail units, anchor store presence or retail mix (Dolega et al., 2016). The introduction of competition into the model was significant, which Huff argued was crucial, as customers typically patronise several stores (Yriogen and Otero, 1983). Thus, the result of the Huff model was catchment areas that were conceptualised as probability surfaces that represented the likelihood of patronage (Dramowicz, 2005), enabling delineation of catchment boundaries, as in a number of studies (Dolega et al., 2016; Davies et al., 2019).

The benefits of the Huff model are numerous, with perhaps the most obvious being that it “filled the gaps”, by addressing competition and accounting for attractiveness in a better way (Wieland, 2017). The model is very flexible, enabling modification of parameters to meet research or business needs (Newing et al., 2015), and is able to simultaneously estimate customer patronage for multiple stores or retail centres (Dolega et al., 2016). In addition, Huff models were shown to generate significant returns on investment for retailers, demonstrating further their effectiveness (Birkin et al., 2010). However, as with the other catchment delineation methods, the Huff model has significant drawbacks. Firstly, when discussing attractiveness, it is not possible to incorporate every factor that might determine this, especially when some of these will be qualitative in nature (Dolega et al., 2016). Secondly, these models are limited by the availability of data from the retailers themselves, their competitors and existing Geographic Information System (GIS hereafter) packages (Segal, 1999), and thirdly there are still problems with the competition element, since a large number of competitor destinations and non-conventional retail formats (e.g., ‘E-commerce’) will create additional complexity (Dolega et al., 2016). Perhaps the most significant limitation is that it is a model, so it is difficult to validate without actual observed consumer behaviours, resulting in great difficulty when trying to capture the dynamicity of choice behaviour, as suggested by Birkin and Heppenstall (2011).

Contemporary catchment techniques

In response to some of the limitations of existing catchment delineation techniques, and a greater availability of high-resolution consumer data (Dramowicz, 2005), new approaches to modelling are emerging (Newing et al., 2015), using such data to provide more accurate delineations of catchments. Such models rely on greater availability of information about customers, typically collected through customer point-of-sale data or loyalty card data – transactional data extracted from the location and purchases made by a consumer either in-store or online (Rains and Longley, 2021). Most interesting with these datasets is their geographic potential, as they contain lots of information about consumers, including where they live and their demographic characteristics (Rose and Dolega, 2022). Thus, using this data in catchment delineations adds significant intelligence (Halsall, 2001), providing a more precise indicator of where customers are coming from (Waddington et al., 2018), enabling derivation of catchments based on observed user behaviours, rather than predicted ones (Davies et al., 2019). As a result, a wealth of research is emerging that uses emerging forms of consumer data to estimate catchments.

Many examples have emerged that use such data and advancements in catchment delineation to generate catchments for individual retail stores. In Waddington et al. (2018), the authors used store trading data from a major UK retailer to explore the links between sales demand and spatiotemporal fluctuations, using data from the loyalty card scheme to cluster stores and develop catchments, with different catchment sizes based on the assigned cluster of the store (Waddington et al., 2018). Similarly, Davies et al. (2019) used data from a UK grocery retailer to generate retail catchments for 95 click-and-collect points across the UK, constructing a bespoke Huff model to generate the catchments. A number of examples have emerged recently which have used similar data to empirically calibrate conventional Huff models; for example, Wang et al. (2016) used social media data from Sina Weibo for Beijing to calibrate a model and estimate catchments for points of interest (POIs hereafter), and Liang et al. (2020) used mobility data from SafeGraph to predict mi.e.,t share (i.e. catchments) for supermarkets and department stores in different U.S. cities.

There are some examples that have used such data and advancements to generate catchments for retail (and town) centres too. For example, in Lloyd and Cheshire (2017) the authors used geo-referenced tweets within the greater London region, to gain greater insights about the patronage of retail centres, constructing catchments based on mobility flows of tweets between the retail centres. This was interesting, as it demonstrated the potential for non-conventional data, such as that from Twitter, in generating catchments. In Jones, Newing and Orford (2022) the authors calibrated a traditional spatial interaction model to estimate catchments for town centres in Wales, accounting for attractiveness, distance and competition more explicitly, integrating expert knowledge about the patronage of town centres in Wales. Furthermore, in Dolega et al. (2016), the authors constructed a bespoke Huff model to delineate catchments for retail centres across the UK, calibrating the model parameters based on data obtained from a survey. However, aside from these three examples, robust empirical catchment delineations for retail centres (as opposed to stores) are sparse in the literature (Pratt et al., 2014), likely relating to the additional considerations needed, as argued by Dolega et al. (2016).

2.5. The *why* – the importance of retail centre geographies

Before introducing *how* retail centre geographies can be used as a geographic data resource in the next section, it is important to discuss *why* such understandings about the retail system and retail centre geographies are important. In particular, some of the major challenges that the global retail sector is facing are introduced, including economic pressures, ‘E-commerce’, changing consumer preferences, and more recently the COVID-19 pandemic and 2022 cost of living crisis, placing significant emphasis on how retail centres are responding to these pressures, to demonstrate the importance of studying these geographical phenomena.

Economic recession

The global economic system can pose significant challenges for the retail sector during times of crisis such as recessions - system-wide shocks that periodically interrupt and disrupt the process of urban growth and development (Martin, 2012), occurring when there are two successive quarters with zero or negative growth (Newson, 2009). The UK experienced recession between 2007-2009, and by mid-2009, the retail sector had been recognised as one of the worst affected sectors; recession affects the retail sector by directly affecting consumers and their shopping habits (Newson, 2009), through reductions in their disposable income, and lack of availability of credit (Thompson, 2013). Consumer confidence fell as customers fearful, keeping money aside for goods that they might need (Newson, 2009; Thompson, 2013). The consequences of these shifts were seen in sales figures; sales flat-lined in response to the economic crisis (Wrigley et al., 2015), particularly for ‘frivolous’ goods (Newson, 2009; Thompson, 2013). As a result, many retailers entered administration and/or liquidation (Tselios et al., 2018); 5,000 stores were affected by closure and/or ceasing of trade (Hutton, 2021), and vacancy rates rose threefold to 12% by 2009 (Tselios et al., 2018).

Given its impacts on retailers and stores, it is no surprise that the economic recession had significant consequences on retail centres; those struggling with high vacancy rates prior to the recession were hit the hardest (Tselios et al., 2018), however the recession also appeared to have significant consequences for the medium-sized retail centres across the UK (Department for Business Innovation and Skills, 2014), which lost out to the larger destination centres across the UK. Significant geographic heterogeneity was seen with these impacts; economic areas in the north performed much worse than those within the south, partially due to their heavier dependence on public sector employment, and higher (pre-recession) vacancy rates in general (Tselios et al., 2018). However, there were very few areas and retail centres that remained completely resilient to the economic crisis, but it must also be noted that other factors (e.g., ‘E-commerce’) were at play during this time.

‘E-commerce’

The internet has altered the daily activities of people in many ways, including the way in which we shop (Weltevreden, 2007). The outcome of this has been the rise of ‘E-commerce’ which has been characterised by significant changes in consumer behaviour and alternative mechanisms through which to shop (Singleton et al., 2016). The sustained growth of ‘E-commerce’ is arguably related to those benefits that it offers the consumer and retailer; for the consumer, the internet offers price comparison, 24/7 convenience, wider variety of products and distribution within a wider reach (Williams, 2009; Wrigley et al., 2015). For the retailer, the internet provides an opportunity to strengthen competitive positioning, particularly for small businesses (Doherty and Ellis-Chadwick, 2010), and offers a better capacity to engage with and understand their customers, through collection of market research data (Basu & Muylle, 2003). Thus, it is no surprise that ‘E-commerce’ has grown; online sales have been growing exponentially in the last 8 years (Singleton et al., 2016), and this growth has continued, with online sales coming to occupy 26.3% of total retail trade in the UK (ONS, 2022)

The rise of ‘E-commerce’ has had significant consequences for physical sites of consumption such as retail centres. Many individual stores have been substituted when failing to offer an online or ‘omni-channel’ option (Singleton et al., 2016; Lansley and Longley, 2017), resulting in the erosion of retailers and a subsequent increase in vacancy rates in retail centres across the UK. Similarly, many retailers have modified their offering, changing the way in which they operate to directly compete with online shopping, as with GAME where they substantially reduced their ‘brick-and-mortar’ offering. However, these impacts have not been homogenous, varying geographically and across the hierarchy of the retail system; large cities are perceived as being immune from the threats of online shopping (Weltevreden, 2007), as more prevalent comparison shopping opportunities make them far more resilient to ‘E-commerce’ (Dixon and Marston, 2002). This trend was particularly evident in Singleton et al. (2016), where the retail centres most susceptible to online shopping were those located in suburban and rural areas of Southeast England and those secondary, medium-sized centres in “clone towns” (Singleton et al., 2016).

Changing consumer preferences (incl. experience economy)

Consumer behaviour has changed significantly in recent years; a greater number of people are looking for opportunities to save time; ‘convenience culture’ (Wrigley et al., 2015), where the expectations of consumers are that they should be able to determine when, where and how they want to shop (Geiger, 2007). Another emerging consumer behaviour is the need for a shopping trip to serve all needs in one location; there are now fewer shopping trips, but the average number of shops being visited per trip is increasing, thus preferences for a multi-purpose shopping experience are growing (Reimers and Clulow, 2009). There are various other demands consumers have of contemporary retail, including authentic shopping experiences that reflect the uniqueness of local communities and are not contrived or formula-driven; the “experience economy”; the modern consumer perceives shopping as an element of leisure, and towns and retail centres need to provide their customers with much more than just retail (Coca-Stefaniak, 2013), thus there needs to be more mixed use of retail space to attract customers in the modern retail environment. Evidence of this trend can be seen in town and retail centres where traditional retail anchors like banks, major clothing stores and public houses are all being replaced by health and beauty services, which all offer an experiential service to consumers (Grimsey, 2018).

The COVID-19 pandemic

The COVID-19 pandemic has caused significant damage to societies, economies, and healthcare systems around the world (Duong et al., 2022). As COVID-19 began to spread around the world in March 2020, countries like the UK saw the imposition of lockdowns and national restrictions, which were implemented by governments to minimise risks to populations and (HSTF, 2021). As a result of these safety measures, the COVID-19 pandemic has and continues to change the ways in which we shop (Sit et al., 2022). The pandemic saw a general shift away from physical 'brick-and-mortar' retailing to online retailing, to prevent possible exposure to the disease. However, in terms of engagement with physical retailing, lots of interesting shifts in consumer behaviour occurred; retail was divided up into 'essential' and 'non-essential', with stores specialising in essential goods (e.g., supermarkets) facing a significant uplift in sales (Nicola et al., 2020; Simbolon and Riyanto, 2020), whilst others (e.g., restaurants) faced significant declines in sales, resulting in large numbers of store closures and redundancies (Nicola et al., 2020).

This has had notable consequences for physical spaces of consumption such as high streets and retail centres. There is now a growing evidence base that the COVID-19 pandemic has accelerated the pre-existing trends of retail decline abysmally, by significantly restricting footfall in many consumption spaces, following the implementation of restrictions to contain the spread of the virus (Enoch et al., 2022). Much of the literature has focused on the consequences of restrictions that were implemented during the earliest stages of the pandemic here in the UK (e.g., national lockdowns). For example, Enoch et al. (2022) identified significant differences in footfall declines between UK town centres during the height of the pandemic. However, a particularly interesting area of research relates these disparities to the characteristics of the retail centres themselves, in particular considering their *functional* role, and *structural* characteristics. In the case of *function*, studies have posited that the economic performance of retail centres during the pandemic was dependent on their overall size and function; many studies have cited that the smaller, localised centres appeared to fare much better than larger towns and cities (HSTF, 2021; Trasberg and Cheshire, 2021;

Enoch et al., 2022; Frago, 2021). In the case of *structural* characteristics, research explored the interactions between retail centre performance and the intrinsic characteristics of the retail centres themselves; Enoch et al. (2022) identified a strong link between increasing vacancy rates and the ability of town centres to recover from the pandemic, and argued that the impacts of the pandemic were a product of the composition (retail, service) and ‘e-resilience’ of retail centres. Similarly, the HSTF (2021) suggested that high streets with a varied retail offer and unique attractions fared much better, and Dolega and Lord (2020) identified a statistically significant relationship with the relative level of deprivation, though not in the context of the pandemic. However, whilst much has been written to unpack and explain the short-term response of centres to the earliest ‘phases’ of the COVID-19 pandemic, much less has been written about the recent ‘phases’ of the pandemic, where the onset of different variants has seen the introduction of different restrictions across the UK. In response to Omicron, the government unveiled a much less stringent set of restrictions than was seen earlier in the pandemic comprising mandatory face masks and vaccine passports (“Plan B”), thus the likely impacts on consumption spaces were reduced in comparison to previous phases of the pandemic, but thus far such insights have not been realised.

Cost of living crisis, energy and 2022

As the world recovers from the economic and social impacts of the COVID-19 pandemic, in 2022 it is being faced with another pressure (Patrick and Pybus, 2022) – the cost-of-living crisis. The world has faced intense pressure from the ongoing war in Ukraine, resulting in a cost-of-living crisis, that has manifested itself into a series of significant rises in energy costs, food prices and taxes, as well as stagnating wages for workers (Patrick and Pybus, 2022). These economic trends are likely to have significant impacts on the retail sector; small increases in interest rates have the potential to destabilise many major retailers, according to Grimsey (2018), thus it is likely that many retailers will be forced into liquidation, as we saw in the 2008 recession. Furthermore, give the rising price of energy and goods, and increased demand for higher wages in line with inflation (Frances-Devine et al., 2022), it is likely that many smaller retailers, who do not have access to a large pool of resources, will not be able to afford to maintain a viable retail business. This is a highly contemporary issues, evolving at a rapid pace, thus there has been no research on how these trends are being manifested in the retail sector, but it is expected to generate significant further decline for retail centres and consumption spaces.

2.6. The *how* – responding to retail sector pressures to protect retail centres and consumption spaces

Thus far some of the major challenges faced by the retail sector in recent years have been outlined, in particular considering the response of retail centres to these pressures. It is vital to empirically quantify these responses and impacts, and identify ways in which to mitigate them. Thus, in this section an overview of some of the approaches to do so is provided; formal recommendations, reports and strategic task forces, before considering the utility of retail centre geographies as geographic data tools for such efforts.

Recommendations and reports

A common approach to understanding the problems facing consumption spaces has been to commission reports by experts in the sector, providing an evidence base upon which to make decisions in government. A notable example was the "Portas Review", undertaken by Mary Portas, or "Mary Queen of Shops" (Cooke, 2018). The aim of the review was to identify how high street decline could be halted; making a number of suggestions as to how best to support the future resilience of high streets (and retail centres). Such suggestions included the creation of small "town teams" to provide a platform through which to discuss high street problems, as well as the strengthening of business improvement districts, making landlords investors of the high street and creating a "national market day" to enable small independent retailers to establish their brands (Portas, 2011). Although Portas' review provided early insights into how the UK retail environment might be revitalised following the 2008 recession, there was strong opposition to her ideas.

One key opposing figure was Bill Grimsey, a renowned retailer and former director of Tesco (Grimsey, 2013). Bill presented the Grimsey review (1 & 2), which offered alternative insights into the problems of British high streets, suggesting ways in which these may be alleviated. Grimsey argued that thus far there had been a failure to assist with struggling high streets; amendments to clause 69 of the localism act in 1988 should have provided local authorities with the mechanism to relieve pressure on retailers, but only a fraction of business rates had actually been alleviated by 2013 (Grimsey, 2013). Grimsey made a number of suggestions, including the creation of a town centre commission, who could undertake high profile reviews of the future of town centres, an open-source network for best practice, the establishment of formal “town centre business plans” and creation of a government high street unit to bridge the gap between central government, local authorities and town centre commissioners (Grimsey, 2013).

The High Streets Task Force

The High Streets Task Force (HSTF hereafter) was an outcome of the Grimsey review; formed to provide hands on support to local areas, by developing innovative strategies to help high streets (and retail centres) evolve and share best practice between them, through assemblage of various stakeholders and experts (HM Government, 2019). The HSTF was launched in 2019 by the Minister for High Streets, and is overseen by the Institute for Place Management (IPM hereafter), operating in a way that varies depending on the place in question; no two places are alike and have their own culture and heritage (HM Government, 2019). However, a common problem facing teams like the HSTF is that they are often subject to financial constraints; thus in 2018 Chancellor of the Exchequer set up a “Future High Streets Fund” which would award up to £25 million for projects seeking to make their high streets fit for the future (HM Government, 2018). As a result, by September 2019, 100 towns had been shortlisted and given £150,000 to develop more detailed proposals.

To support the decisions made by governments, the IPM have conducted large volumes of research about retail centres and high streets to provide a strong evidence base for such decisions. These studies have included discussions about problems facing high streets based on observations and data (Parker et al., 2017; Millington et al., 2015), evaluating the successes and failures of Business Improvement Districts (Grail, 2015), developing high resolution classifications of high streets (Mumford et al., 2021), providing evidence of the performance of consumption spaces throughout the pandemic (Mumford et al., 2021), and discussing strategies for post-pandemic recovery (Ntounis et al., 2020). Such insights are vital, as they are able to provide robust and timely evidence to support decision-making, utilising the expertise of academics and practitioners in the field.

Retail centre geographies as a geographic data resource

Given the increasing availability of data, and vast network of literature documenting the *who, what* and *where* of retail centres, research is rapidly emerging that seeks to use retail centre geographies as geographic data resources to better understand some of the problems facing high streets, retail centres and consumption spaces. In particular, a large body of work has been conducted by researchers at the University of Liverpool and CDRC on this topic. For example, given one of the biggest challenges facing physical retailing and retail centres is ‘E-commerce’, it is no surprise that efforts have also been made to use retail centre geographies as data tools to understand its impacts. In their paper, Singleton et al. (2016) proposed a mechanism through which to assess the impacts of online shopping on retail centres in the UK, by integrating elements of both demand (Internet User Classification) and supply (‘high-risk’ retail categories), the authors were able to derive an ‘e-resilience’ score for retail centres, which quantified their vulnerability or resilience to online shopping. The measure enabled insights about the geographies of ‘e-resilience’, and contributed significantly to better understanding the consequences of online shopping on retail centres, demonstrating the importance of retail centre definitions, as without understanding of the *where, what* and *who* of UK retail centres, such insights would not have been possible.

Similarly, Dolega and Lord (2020) utilised the same retail centre definitions to examine their economic performance over time, for Liverpool City Region. Examining changes to vacancies within different retail centres, the authors were able to unpack the geography of retail success and decline, identifying the characteristics and attributes of retail centres that are likely to have a viable future (Dolega and Lord, 2020). The importance of this research was in highlighting that retail success and decline is not geographically homogenous, thus further demonstrating the value in national level understandings of retail centre geographies, especially when integrated with other ancillary information. These findings echo the work of Jones, Newing and Orford (2022), who constructed a typology (the *what*) and catchments (the *who*) for Welsh town centres and using vacancy data from LDC, were able to investigate the economic performance of town centres in Wales. Also, in response to the COVID-19 pandemic, Trasberg and Cheshire (2021) utilised the CDRC retail centre definitions to ask whether retail centres specialising in certain types of goods or with specific functions experienced different levels of decline following the onset of the pandemic. Their research was of great significance, providing evidence that suburban high streets were the most resilient, whilst those specialising in premium goods experienced the greatest declines (Trasberg and Cheshire, 2021).

Retail centre geographies are thus of great importance when it comes to understanding the causes and consequences of retail decline. As a geographic data resource, comprising information about *where* they are located, *what* characteristics they have and *who* uses them (i.e., catchments), retail centre geographies can contribute significantly by providing a geographic data resource that can be used to show *how* the retail sector is responding to these external pressures, providing a formal evidence base upon which to make better evidence-led decisions. Furthermore, this demonstrates evidence for the for the *why;* why retail centre geographies are needed to generate such insights.

3. A framework for delineating the scale, extent and characteristics of American retail centre agglomerations

The content of this chapter is published as a research paper in Environment and Planning B: Urban Analytics and City Science:

Ballantyne, P., Singleton, A., Dolega, L., Credit, K. 2022. A framework for delineating the scale, extent and characteristics of American retail centre agglomerations. *Environment and Planning B: Urban Analytics and City Science.* 49 (3), 1112-1128.

DOI: https://doi.org/10.1177/23998083211040519.

Chapter Overview

This chapter represents the first of three empirical chapters in this thesis, and fulfils the first aim of this PhD thesis by developing an analytical framework through which to generate retail centre geographies in new international settings. In particular, it evaluates the potential for innovations in relevant methodologies (e.g., Pavlis et al., 2018; Dolega et al., 2021) to generate such insights, focusing specifically on the Chicago Metropolitan Statistical Area (MSA hereafter) as a ‘pilot’. It integrates hierarchical density-based spatial clustering of applications with noise (HDBSCAN hereafter) with network distance matrices and H3 to delineate boundaries, and constructs a multidimensional typology using the framework proposed by Dolega et al. (2021), to generate understandings about the *where* and *what* of retail centres in Chicago MSA. In addition, the study provides early evidence of *how* retail centres can be used to understand the impacts of COVID-19, through examination of trends in consumer visits between different groups of retail centres, during the early weeks of the pandemic.

The key contributions of this chapter are as follows. a set of retail centre definitions and accompanying typology are presented for a new geographic setting, representing the first empirical attempt to do so. Secondly, the potential for recent methodological innovations to generate such understandings at the national level is evaluated, concluding that further conceptualisation and methodological tools are needed. Thirdly, early evidence of the impact of the COVID-19 pandemic on retail centres is provided, demonstrating that these impacts are profound, and that future research should utilise retail centre geographies as tools through which to further investigate these impacts.

Abstract

Retail centres are important tools for understanding the distribution and evolution of the retail sector at varying geographical scales. This paper presents a framework through which formal definitions and typologies of retail centres, such as those in the UK, can be extended to the US. Using Chicago as a case study and data from SafeGraph, a retail centre delineation method is presented that combines HDBSCAN with ‘H3’, and demonstrate the usefulness of a non-hierarchical approach to retail classification. In addition, the dynamicity and comprehensibility of retail centres is demonstrated through their use as an effective tool through which to better understand the impacts of COVID-19 on retail centre ‘health’, demonstrating significant scope for a comprehensive delineation of the scale, extent and characteristics of American retail centre agglomerations, providing a tool through which to monitor the evolution of American retail.

3.1. Introduction

The contemporary physical retail environments of cities and urban areas have complex form and function, evolving in response to a multiplex of pressures. In the UK, the effects of rising online sales and the 2008 recession continue to be felt on high streets, where retail presence continues to decline (Dolega and Lord, 2020). Similarly, the American retail sector is amid an ‘apocalypse’ (Boerschinger, Pansch and Lupini, 2017; Isidore, 2017) with notable decreasing ‘brick-and-mortar’ sales and high vacancy rates (Boerschinger et al., 2017), and more recently, the global impact of the COVID-19 pandemic have been particularly visible, exacerbating many of these problems (Nicola et al., 2020).

Understanding better the dynamics of retail evolution - occurring partially in response to these pressures - is vital for academics and various stakeholders, requiring attention at a range of spatio-temporal scales and intensities (Thurstain-Goodwin and Unwin, 2000). In particular, having ways to monitor the ‘health’ of retail agglomerations has become acutely important, given their role in economies and communities (Berman and Evans, 2013; Coca-Stefaniak, 2013). However, in order to monitor the health of retail centres it is first vital to understand their form and function. In particular, emphasising the development of automated approaches to estimate their spatial extent, monitor evolutionary trajectories and derive catchment characteristics (Joseph and Kuby, 2016), providing stakeholders with a platform to make evidence-led decisions.

As such, the contributions of this paper are threefold. Existing frameworks are built upon to provide a theoretical rationale and empirically grounded framework for the i) definition and ii) characterisation of retail centres, using innovative methods and new forms of data, before demonstrating its application to the Chicago MSA, and iii) using it to highlight the visible impacts of COVID-19, for the first time, on the ‘health’ of different retail centres. Chicago is an interesting setting to implement this framework, given the existing wealth of research on urban retail structures (Berry, 1963; Casparis, 1969; Joseph and Kuby, 2016; McMillen, 2003), and significant gap in the use of contemporary methods and data to improve such understandings.

3.2. The *'Place'* of Retailing

The tendency for retail units to agglomerate has received substantial theoretical attention. Often being linked to the economic decisions of individual businesses (Sohn et al., 2003), which concentrate spatially to benefit from 'agglomeration economies' (McCann and Folta, 2008). Thus, given the economic advantages of spatial clustering, it is perhaps no surprise that centrality and agglomeration are considered key concepts in the geographies of retail space (Brown, 1992). A number of theories and models have been posited to conceptualise these geographical tendencies. Christaller's CPT is widely regarded as a key model, proposing that 'central places' exist to serve the need for goods in surrounding areas (Parr, 2017). At a relatively simple level, CPT provides a conceptual model to better understand the spatial arrangement of retail (centres). However, CPT has been criticised for unrealistic assumptions about consumer behaviour (Parr, 2017), and it fails to apply in polycentric cities or irregular commercial forms (Brown, 1992; Dolega et al., 2021). Other conceptualisations include the theories of Von Thunen and Haig, but both have faced similar criticism (Brown, 1992; O'Kelly and Bryan, 1996). Although there is scope in applying these principles, they are limited in failing to address the complexity of structural and functional interdependencies between centres (Dolega et al., 2021).

The Spatial Extent of Retail Agglomerations

Historically, there have been numerous attempts to differentiate retail agglomerations based on form (e.g. Proudfoot, 1937). The first form-function delineation (Berry, 1963), identified types of retail clusters in Chicago, quickly becoming the universally accepted model of retail organisation. Furthermore, Murphy and Vance provided a delineation utilising a 'central business index' (Murphy and Vance, 1954), and Brown (1992) proposed the first 'non-hierarchical' form-function delineation. More recently, the coupling of new data and analytical frameworks have revived interest in delineation (Dolega et al., 2016). Early "data-intensive" work (e.g. Thurstain-Goodwin and Unwin, 2000) utilised continuous density transformations to delineate UK town centres, and in other examples surface density functions coupled with volume contours and geometric operations have been used (Singleton et al., 2011; de Smith et al., 2018). Furthermore, Ordnance Survey and Geolytix have constructed similar retail centre definitions, but were limited in exclusion of some retail functions and lack of open accessibility respectively.

In recent examples, there has been preference for more explicit definitions using spatial cluster analysis of store locations (Yoshimura et al., 2021; Han et al., 2019; Lloyd and Cheshire, 2017). For example, Pavlis et al. (2018) utilised an unsupervised machine-learning algorithm (DBSCAN) in automated delineation of UK retail centres., developing a modified version of the algorithm to solve a common issue with the application of DBSCAN to real world distributions; heterogeneities in local point density (Campello et al., 2013). However, to facilitate issues with the dataset, their approach required specification of additional parameters, and was reasonably computationally inefficient, limiting its implementation in future studies.

The HDBSCAN algorithm has developed saliency, offering a solution to many of DBSCAN's limitations. With only one mandatory input parameter (*minPts*), HDBSCAN makes parameter selection more intuitive and robust, whilst accounting for heterogeneous point densities, through production of a DBSCAN cluster tree (Campello et al., 2013). Furthermore, the algorithm can utilise precomputed distance matrices for improved performance (Campello et al., 2013), enabling incorporation of network distances. Despite this, its potential for retail centre delineation has not yet been realised.

The Typologies of Retail Agglomerations

Classifications of retail agglomerations have traditionally argued that retail is hierarchically organised (Brown, 1992; Dolega et al., 2021), and can be classified based on assumptions about demand and supply, drawn from CPT. These 'vertical' classifications have been criticised for using simple datasets and non-uniform methods, failing to accurately represent the spatiality of retail provision (Dolega et al., 2021; Guy, 1998). Recently there has been a call for classifications that better comprehend changes to retail provision (Grewala et al., 2017), which are now both possible and more necessary (Dolega et al., 2021), including the use of a socioeconomic classification matrix (Coca-Stefaniak, 2013) and footfall patterns (Mumford et al., 2017).

In a recent example, using the boundaries delineated in Pavlis et al. (2018), Dolega et al. (2021) used a data-driven approach to construct a 'non-hierarchical' typology. Variables were gathered to capture four domains they believed key to capturing the multidimensionality of retail centres, and using an unsupervised machine learning algorithm called partition around (the) medoids (PAM hereafter), they developed a two-tier classification, with PAM used over k-means to reduce the impacts of outliers (Struyf et al., 1997). This approach arguably provided the most nuanced and comprehensive way of representing relationships between centres, rather than assuming that hierarchical relationships prevail. However, in the US systematic nationwide and rigorous “data-intensive” studies on the scale, extent and characteristics of retail agglomerations are yet to be realised.

3.3. Study Context and Data

The Chicago MSA has a retail sector that provides a rich consumption experience for residents and visitors (Glaeser et al., 2001). However, as across the US, the sector remains challenged by increasing retail vacancies (Joseph and Kuby, 2016), an over-saturation of “brick-and-mortar” retail, and increased uptake of ‘E-commerce’, resulting in shifting shopping habits and store typologies across Chicago (Joseph and Kuby, 2016). Thus, there is justification for a contemporary definition of retail centres, which can shed insight into current retail provision. However, a challenge in defining centres is a lack of comprehensive, up-to-date and open-access retail location data. In this study, data from SafeGraph is used as the best available source, in particular their register of ‘core places’ where consumers spend money or time in the US (SafeGraph Inc., 2020a), and corresponding mobility data or ‘weekly patterns’ (SafeGraph Inc., 2020b), collected from the global positioning system (GPS hereafter) data of 45 million anonymised mobile phone users (Gao et al., 2020). SafeGraph ‘core places’ were re-classified to identify the ‘retail places’, and the ‘non-retail’ places were removed from the dataset, leaving 106,058 retail locations for the Chicago MSA. For background information on the datasets and processing of these see Appendix I.

3.4. Delineating Urban Retail Centres

HDBSCAN was adopted to derive retail centres for the Chicago MSA (Figure 1), using the retail places extracted in Section 3.3. As above, *minPts* is the only mandatory parameter in HDBSCAN, controlling the minimum number of points in a cluster. The value was set to ten to maintain a consistent definition with Pavlis et al. (2018). Network rather than euclidean distances were used in HDBSCAN - the lengths of the shortest path (by road) between points - to better account for the role of urban morphology in retail distributions (see Appendix I). HDBSCAN was iterated for subsets of points delimited by each county to enable practical run times for generation of the network distance matrices (*dist)*. However, in the case of ‘Cook

County', the largest in the MSA, HDBSCAN was iterated for three subsets naturally delimited by the Chicago River. These iterations had little effect on centre distributions, but were deemed inevitable as computationally this limitation could not be avoided. As a result, HDBSCAN generated cluster ID's for every point within a significant cluster, labelling other points as 'noise'.

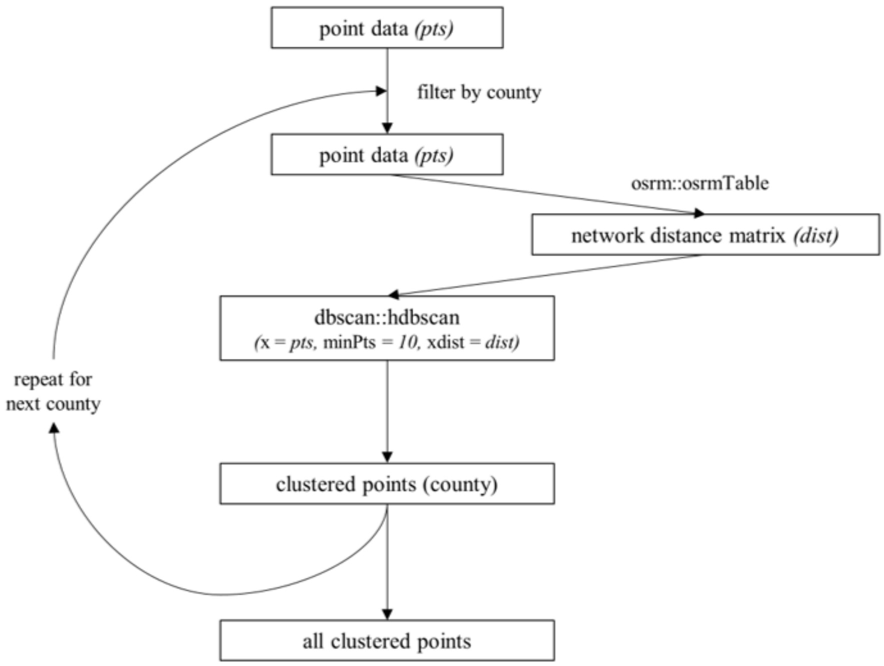

Figure 1. Iterative application of network-based HDBSCAN for delineation of retail centres.

The output of HDBSCAN are points with cluster labels, thus not complete demarcations of areas, and required refining to remove points outside the main cluster 'core' (Figure 2). Thus, an approach to derive and refine boundaries utilising the 'H3' hexagonal spatial indexing system was developed (Uber, 2018), seen below in Figure 2. Using the 'h3jsr' R package (O'Brien, 2020), each clustered point (2A) was aggregated to a hexagon at resolution 11 (2B), each having an area of approximately 10 metres. A buffer consisting of the six neighbouring hexagons (k-ring) was extracted (2C). Using the *minPts* threshold defined in HDBSCAN, only those contiguous and non-isolated zones containing ten or more retail places were extracted as the final retail centres (2D). The assigned cluster ID's were comprised of a county and numeric identifier, with clusters in Cook containing additional identifiers – W, S, N – to reflect the intra-county iterations.

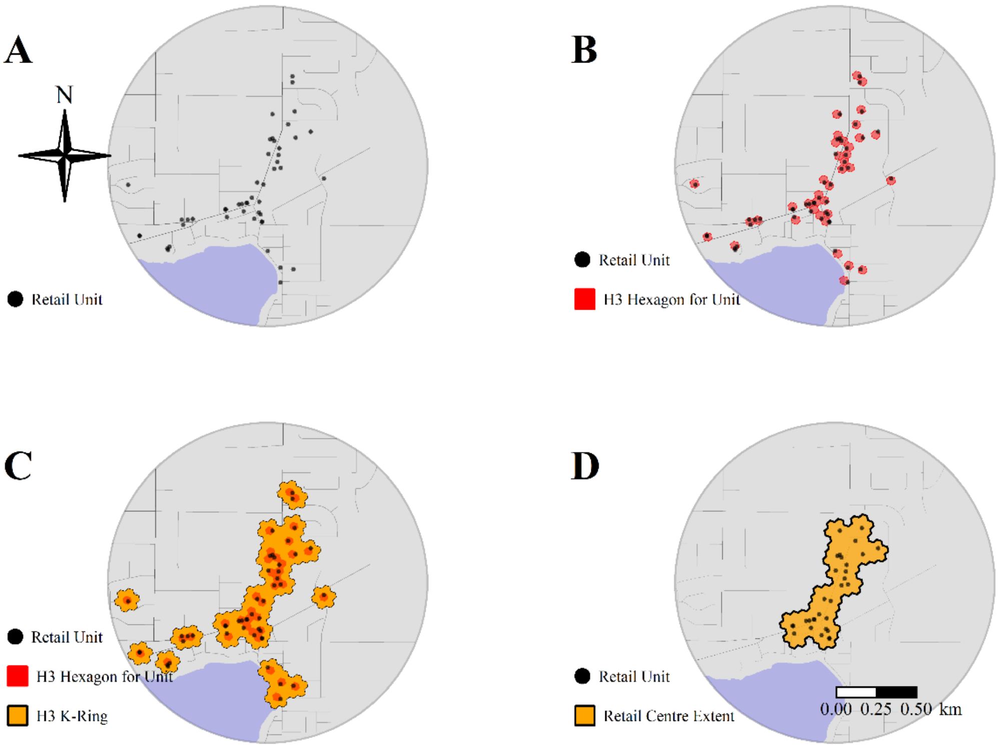

Figure 2. Cluster refinement using H3.

The Spatial Extent of Retail Centres

This approach extracted 1,599 retail centres, with the smallest composed of 10 units (*minPts*) and the largest; 'The Loop' (SC 1) 2,013 units. The majority were located in the CBD and nearby suburbs; the areas of greatest economic activity (Pan et al., 2017; Sohn et al., 2003). Unsurprisingly given its polycentricity, many agglomerations also existed within the core of cities in the wider MSA like Elgin and Joliet (McMillen, 2003), and along major transport arteries (McMillen, 2003; Pan et al., 2017; Sohn et al., 2003). To assess the effectiveness of the delineation, two case study areas were chosen based on relevant literature and to highlight the efficacy of the algorithm in different urban settings – Chicago CBD and Schaumburg Village.

In Chicago CBD (Figure 3), one large cluster was identified (SC1), unique in terms of size and morphology, representing the 'historic retail core' of Chicago; The Loop (Credit, 2020). NC1 encompassed the 'Magnificent Mile' (Figure 3), another significant retail destination (The Magnificent Mile Association, 2015), and WC2 corresponded to Fulton Market. Schaumburg Village also had a large concentration of retail centres (Figure 4), unsurprising given its reputation as a ‘golden retail corridor’ (Fleming, 2009). Major shopping developments were delineated, such as Woodfield Mall (WC3) and Woodfield Green (WC27), as well as some smaller centres. This was arguably only possible through integrating network distances, the most effective way to understand Chicago’s urban structure (Pan et al., 2018). However, the use of building geometries over points would arguably generate more accurate centre boundaries (OS, 2019) by fully accounting for the wider footprint of retail locations (e.g., shopping malls).

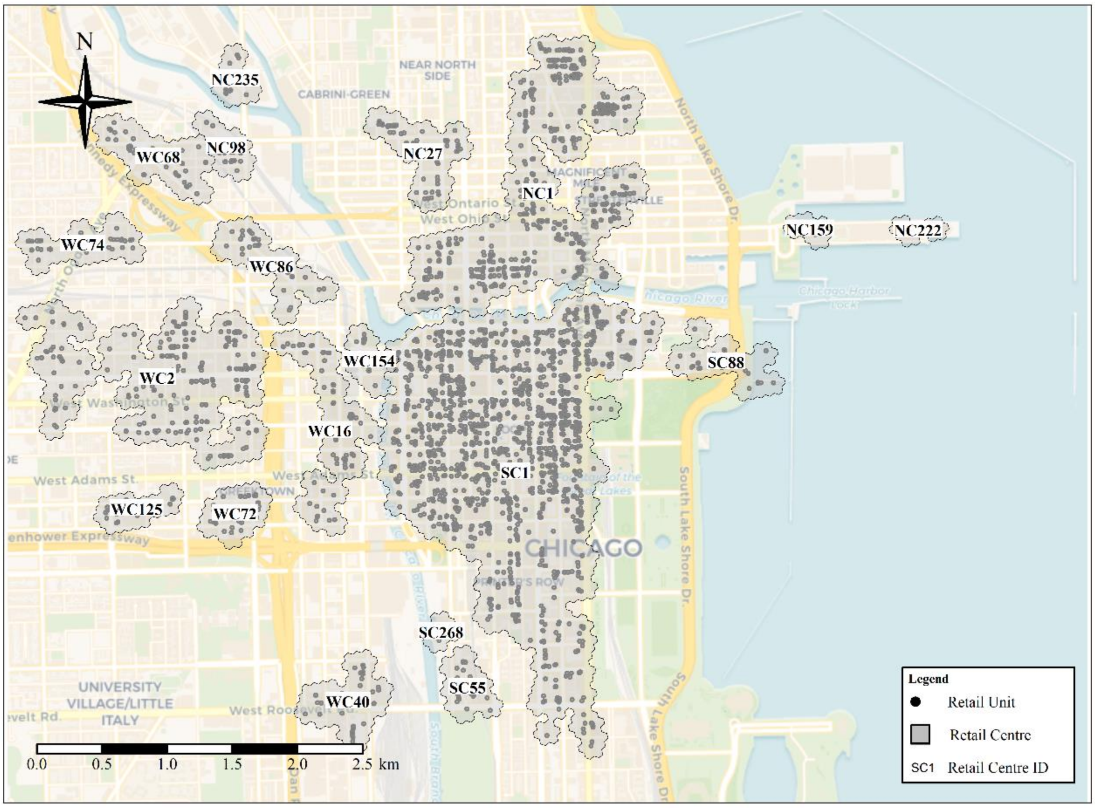

Figure 3. Delineated retail centres for Chicago CBD.

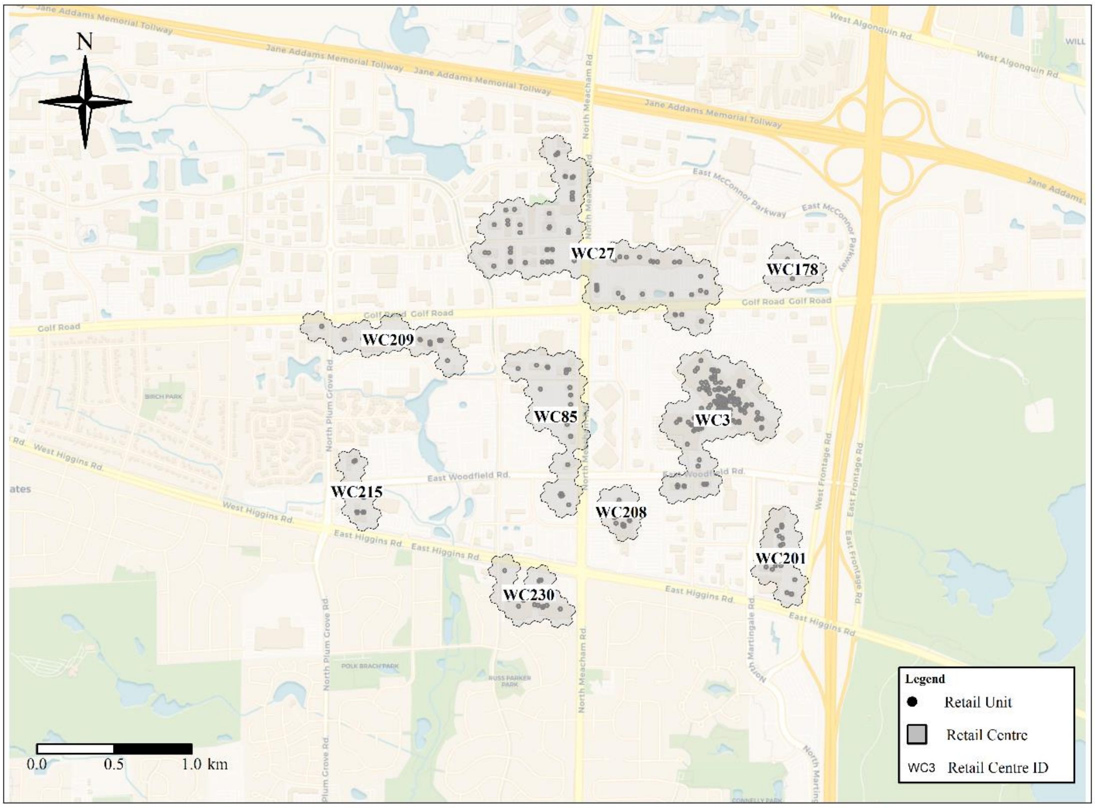

Figure 4. Delineated retail centres for Schaumburg Village.

To consider the validity of the retail centre boundaries, they were compared with other spatial data on retail distribution - SafeGraph 'patterns' (SafeGraph Inc., 2020b) and employment data from the US Census Bureau (2018). In the example below for Chicago CBD (Figure 5), the centre boundaries seemed to align closely with the spatial 'signature' created by the 'patterns' data (5A) and encompassed the majority of census blocks identified as having a high proportion of retail employment (5B), with those not encompassed being sites of small retail centres (< 10 units). Overall then, based on the authors collective understanding of the region and quantitative validation of the retail centres (Figure 5); this approach has arguably identified a set of retail centres that robustly summarise the structure of retail in metropolitan Chicago. Retail boundaries, especially in large urban areas, could be challenged based on the public perception; however, such an empirical delineation has clear advantages including the ability to updated over time, something not feasible with perceptions. Interestingly, the centres themselves vary in location, scale and extent; therefore, in order to derive a nuanced picture of their multidimensional characteristics and position within a system of (metropolitan) retail, these sites are next explored from a typological perspective.

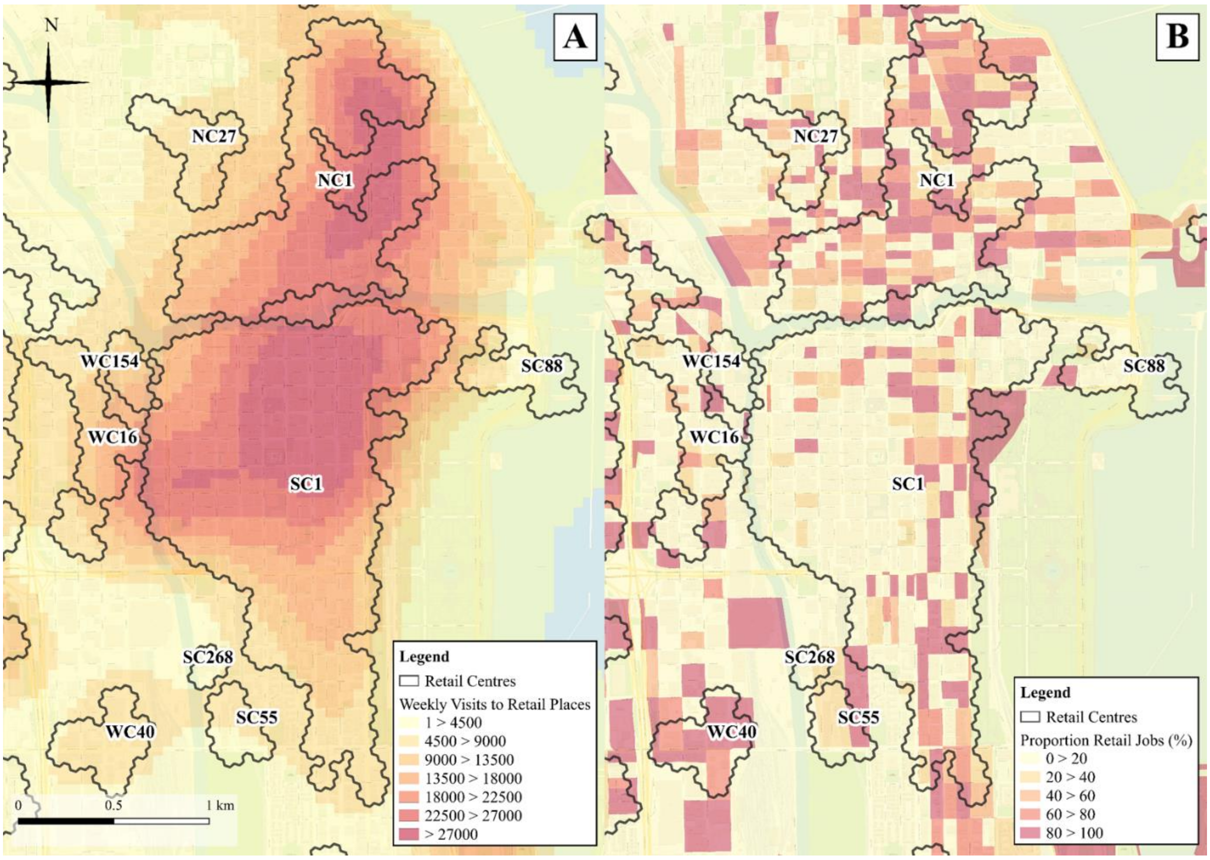

Figure 5. Validation of retail centre boundaries using SafeGraph 'patterns' (A) and data from the US Census Bureau (B).

3.5. A Typological Perspective on Retail Agglomerations within Chicago MSA

To develop a comprehensive classification for the retail centres, this study adopted the methodological framework developed by Dolega et al. (2021). Twenty-four variables were selected to align with those used for each domain in Dolega et al. (Table 1), with the vast majority derived from the retail locations themselves. However, in order to account for 'economic health', the 'weekly patterns' dataset was used (SafeGraph Inc., 2020b), with the proxy variables (visits, dwell and distance travelled) used over vacancy rates and/or level of online exposure (as in Dolega et al., 2021), as the latter were not available. Principal component analysis (PCA hereafter) was performed (Mumford et al., 2017), revealing significant variation in all variables, but four were removed due to issues with multicollinearity and a lack of coverage in Chicago. The remaining twenty variables can be seen below in Table 1, with more detail on each of the variables found in Appendix I (sub-section C).

Table 1. The retail centre classification framework (Dolega et al., 2021), and variables used to implement the framework in this study.

Domain	Domain Description	Variables
Composition	Classifying retail centres by the types of store present and purposes of shopping trip	Proportion of comparison, convenience, service and leisure units in each centre
Diversity	Focusing on the variety of goods and services offered, and the variety of ownership of stores in each centre	Proportion of independent retail units, diversity of SafeGraph 'top-categories'
Size & Function	Identifying the various roles of retail centres and the ways in which they interact with catchment geodemographics	No. of units, linearity (roeck score), median distance travelled, proportion of catchment population occupied by geodemographic groups
Economic Health	Exploring economic performance of retail centres by measuring the drivers of its vitality and viability	Median dwell time, median weekly visits

Before running the classification, the variables were standardised and the optimal *k* value was determined using a clustergram in conjunction with average silhouette scores, to counteract any subjectivity in clustergram interpretation. The classification – using PAM - was performed twice, extracting a set of five retail centre ‘groups’ and ten nested ‘types’. The utility of this classification was enhanced by providing additional descriptive profiles highlighting their salient characteristics, summarised below in Table 2.

The Geography of Retail Centre Characteristics

The first group of centres typically existed at the ‘core’ of urban areas like Chicago and Elgin, and along established retail strips like Ogden Avenue. Inner city leisure (3.1) was concentrated in the CBD, whilst suburban leisure (groups 3.2 and 3.3) was more geographically dispersed. The distribution of comparison centres was also uneven, with the leading destinations (2.1) typically found in well-established retailing developments (e.g. “Fashion Outlets of Chicago”). As suggested by Casparis (1969), convenience and service retail (groups 4 and 5) was dispersed throughout the MSA, with secondary convenience (4.2) and service centres (5.2) concentrated in urban centres, whilst primary centres (4.1, 5.1) were found in suburban neighbourhoods.

Table 2. Salient characteristics of retail centre groups and nested types, including summary of their 'health' in response to the COVID-19 pandemic.

Supergroup	Group	Key Characteristics	COVID-19 'Health'
1. Large Multipurpose Centres & Historic Retail Cores	1.1 Large Multipurpose Centres & Historic Retail Cores	Large no. of visits, diverse retail offering with higher proportions of service than other retail types	Decreased following 'Stay at Home' order, sustained reduction until September 2020
2. Popular Comparison Destinations	2.1 Leading Comparison Destinations	Greater visit frequency, diverse offering, serving neighbourhoods of "Wealthy Nuclear Families" and "Old Wealthy White"	Slight increase following 'Stay at Home' order, sustained until September 2020
	2.2 Secondary Comparison Destinations	Lower visit frequency, less diverse offering, larger number of chain retailers, serving "Middle Income Families", "Low Income Families" and Hispanic and Kids"	
3. Leisure Strips	3.1 Inner City Leisure	Highest proportions of leisure and independents, concentrated in "Wealthy Urbanite" neighbourhoods	Slight decrease following 'Stay at Home order, sustained until September 2020
	3.2 Popular Suburban Leisure Centres	High proportions of leisure and independents, longer dwell times, "Middle Income" and "African-American Adversity" neighbourhoods	
	3.3 Secondary Suburban Leisure Centres	Compact, higher proportions of chain leisure-based retail, concentrated in "Low-Income" and "Hispanics and Kids" neighbourhoods	
4. Independent Service Centres	4.1 Diverse Service Centres in Affluent Neighbourhoods	Highest proportions of services, compact, large number of independents, serving neighbourhoods of "Wealthy Nuclear Families" and "Old Wealthy White"	No change following 'Stay at Home' order or later into the year

	4.2 Service Centres in Hispanic and Low-Income Neighbourhoods	Large number of independents, longer dwell times, concentrated in "Hispanic & Kids", "Low Income" and "African-American Adversity" neighbourhoods	
5. Small, Local Convenience Centres	5.1 Primary Convenience Centres	Greater visits, compact, dominant in neighbourhoods of "Wealthy Nuclear Families", but also "Old Wealthy White"	Slight increase following 'Stay at Home' order, sustained until July, back to pre-COVID-19 levels by September 2020
	5.2 Secondary Convenience Centres	Smaller and more linear, less visits, less chain retailers, longer dwell times, found in many different geodemographic neighbourhoods	

In the case of Chicago CBD (Figure 6), it was unsurprising that both SC1 and NC1 were identified as being group 1 centres, given their status as major retail corridors (Credit, 2020; Magnificent Mile Association, 2015). Furthermore, the density of leisure-based centres in the CBD was also expected, as a major hub for bars and restaurants. Schaumburg Village had a notable concentration of primary comparison destinations (Figure 7), such as WC3 (Woodfield Mall), but also one secondary comparison centre (WC201) with a characteristically smaller and homogenous comparison offering. The overall dominance of comparison and service centres in Schaumburg is arguably unsurprising, as these sectors provide a significant majority of local employment (McMillen, 2003). Thus, as demonstrated here, this approach to retail classification arguably provides an accurate and robust representation of the structure of retail across Chicago MSA.

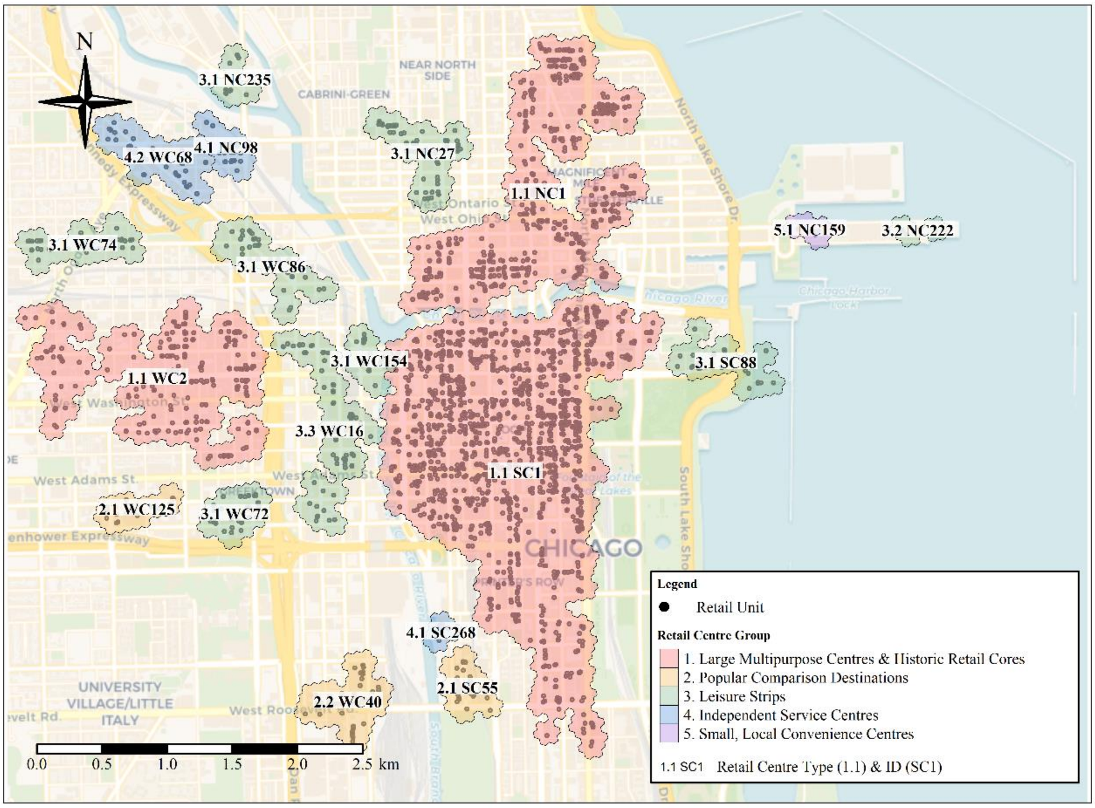

Figure 6. Delineated and classified retail centres in Chicago CBD.

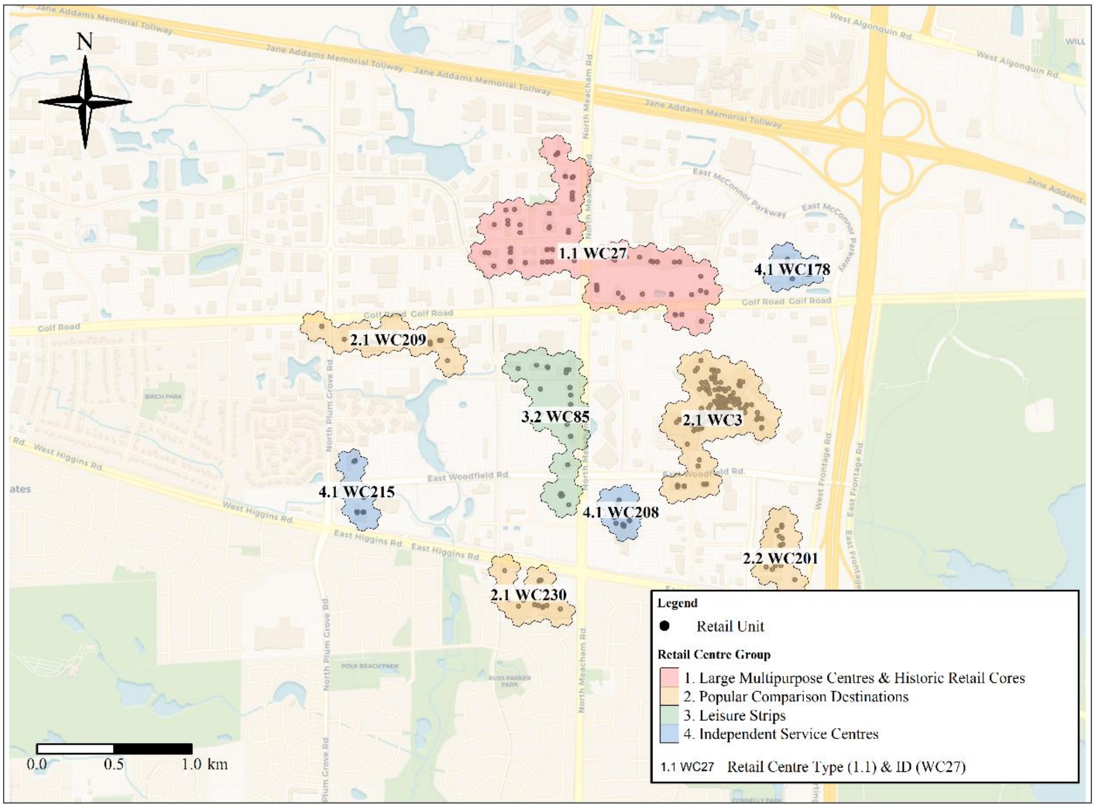

Figure 7. Delineated and classified retail centres in Schaumburg Village.

3.6. COVID-19: Demonstrating the Utility of Retail Centres

A plethora of studies have used retail centre definitions and their typologies to understand wider retail sector processes (e.g. Singleton et al. 2016; Lloyd and Cheshire, 2017), but with the exception of AbedRabbo et al. (2021), their application to understanding the COVID-19 pandemic has been limited. It has been widely documented that the pandemic has exacerbated sector challenges through enforced restrictions on retailers (e.g. 'Stay at Home' orders). Interestingly however, some retailers have faced greater challenges than others (Nicola et al., 2020), creating indirect disparities in the 'economic health' of traditional retail agglomerations, a trend that has not yet been quantified in relation to retail centres. In this final section, adopting an exploratory approach, an immediate use-case for the centres delineated in this study (and typology) is demonstrated, through exploration of changes in visits to centres, as a proxy indicator of their 'health' (Bonaccorsi et al., 2020). Using the SafeGraph 'weekly patterns' dataset (Appendix I), disparities in the effects of COVID-19 on the 'health' of different structures and functions of retail in Chicago are quantified, through exploration of visits to the retail centres over a 12-month period, contributing to existing literature utilising similar mobility datasets (Bonaccorsi et al., 2020; Gao et al., 2020).

Retail Centre Dynamics and COVID-19

In general, the retail sector saw significant decreases in overall 'health' (Figure 8) coinciding with the 'Stay at Home' order (Pritzker, 2020), contracting in total visits by 1/3 in one week, and remaining suppressed until the end of April. Following Pritzker's announcement, many retail centre groups saw decreased visit share, most notably the first group of centres, where share was down 2% (Figure 8). In contrast, the "small, local convenience centres", saw increases in share that were sustained throughout April and May. This trend suggests a general shift from large city centre agglomerations towards the smaller, more local ones, typically offering greater proportions of 'essential goods' and performing better in terms of 'economic health' (Roggeveen and Sethuraman, 2020). What is surprising is that comparison centres did not seem to exhibit any notable decreases in 'health', despite documented declining popularity in 'non-essential' goods (Roggeveen and Sethuraman, 2020).

In the longer term, visits to retail centres remained suppressed, with interesting implications for the long-term 'economic health' of centres. Visit share around group four and five centres appeared to be returning to pre-COVID-19 levels, but group one centres continued to occupy a 4-5% reduction on average in visits, suggesting consumers continued to visit "primary comparison destinations" over the more 'traditional' shopping locations. This is interesting and could contribute to speculation that traditional high streets are facing accelerated decline as a result of the pandemic, therefore potentially becoming no longer 'fit for purpose' as retail distribution networks. Whilst acknowledging the complexity of these processes, and the need for advanced modelling techniques to better quantify them, it is argued that this approach and its findings are timely and significant. The insights generated into the apparent impacts of COVID-19 on the 'economic health' of different retail structures/functions in Chicago are useful and novel. Furthermore, the utility of the retail centre framework proposed in this paper is demonstrated, by using it to contribute to a growing evidence base in a rapidly emerging field of research in retail – COVID-19.

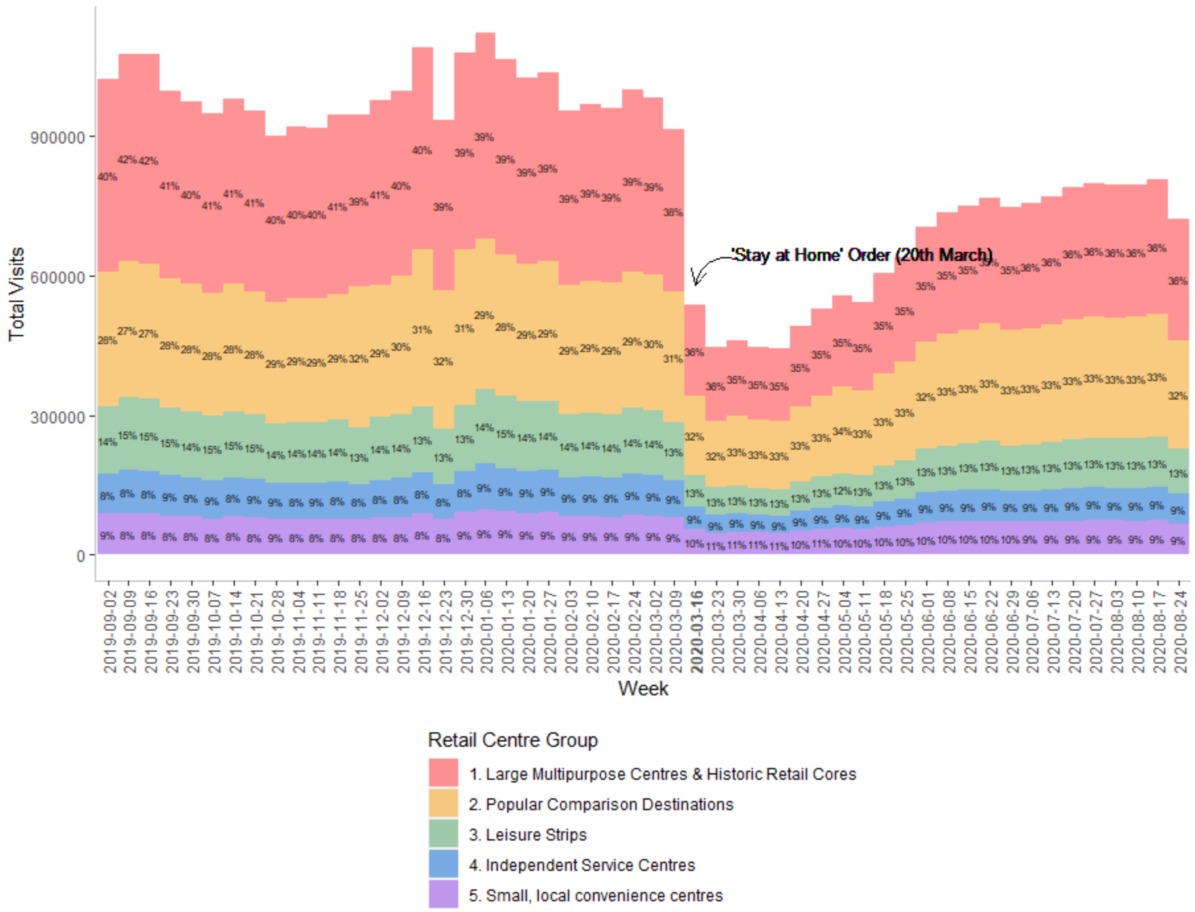

Figure 8. Change in total retail centre visits over the 12-month period, disaggregated by retail centre group to illustrate change in share (%).

3.7. Discussion and Conclusions

This study has enhanced and extended, to an American setting, a data-driven framework for the derivation of retail centre agglomerations, specifically for the Chicago MSA. Using data from SafeGraph, retail centres were delineated through integration of HDBSCAN and 'H3', and the functional ecologies of the 1,599 retail centres were presented as a 'two-tier' classification, constructed using the PAM algorithm. Finally, an immediate use-case for the framework and its outputs (retail centres) is given, in providing insights as to the role of the COVID-19 pandemic on the 'health' of different structures and functions of retail.

Methodologically, this paper has demonstrated the effectiveness of HDBSCAN as a simpler and faster alternative to the modified-DBSCAN approach used in Pavlis et al. (2018). This arguably makes future delineations within other international settings more feasible, however there are scalability concerns when accounting for street networks. The classification framework used here is also of significance; demonstrating its first international application since its conception (Dolega et al., 2021). Using variables deemed fundamental to understanding the contemporary retail landscape, and classification based on similarity and salient characteristics, a more representative insight into the spatiality of retailing has been provided (Dolega et al., 2021; Guy, 1998), than has been produced by other hierarchical or non-hierarchical classifications, both in and out of Chicago (Coca-Stefaniak, 2013; Brown, 1992).

Comprehensive retail centre definitions such as this have significant implications, contributing valuable insights into the interplay between external pressures and physical retail space, through indirect assessment of their evolutionary trajectories. Such insights can also contribute to the academic rigour on 'E-commerce' in the US (Grewala et al., 2017), for example through a greater understanding of the geographies of internet usage, a direct quantification of the 'resilience' of American retail centres to 'E-commerce' could be constructed (as in Singleton et al., 2016). Most interestingly however, is the evident need for additional research to unpack the complex relationship between the retail sector and the COVID-19 pandemic, utilising this framework and its outputs to provide a stronger understanding of the wider retail sector response, not just specific store types. In particular, there is significant potential in the modelling of various retail centre attributes (e.g. diversity, catchment geodemographics) in relation to 'economic health', as defined by metrics (e.g. Comber et al., 2020) rather than proxy indicators, to better comprehend the role of COVID-19 on the evolving American retail landscape. It is however apparent, that in order to achieve such insights, there must first be an understanding of *where* these retail agglomerations are, *what* characteristics they have, and *who* is using them.

On this basis, it is proposed that there is significant scope for a delineation of the spatial extent and characteristics of retail centre agglomerations for the national extent of the US. Future research is needed to ensure a computationally more scalable approach to retail centre delineation, that is not limited to metropolitan areas. Such an increase in scale would also enhance the resolution of an American retail centre typology, through greater abundance and variance in centres (and characteristics) and incorporation of specific niches in American retail and urban morphology. It is also important to acknowledge that the approach and outputs are heavily influenced by the input retailer location data. However, the SafeGraph 'core places' provides the most comprehensive, up-to-date and openly-accessible register of businesses in the US, and as such has significant potential in a proposed geographical expansion of this research. Such an expansion, utilising the framework and dataset posited here, would generate substantive insights into the spatiality of local, regional and national retail provision, whilst also providing a set of tools through which to better understand how retail provision continues to transform. Looking forward, this will be essential as American retail continues to traverse the COVID-19 pandemic and 'retail apocalypse'.

Supplementary Materials

To see the accompanying supplementary materials for this chapter, please see Appendix I.

Acknowledgements

The authors would like to thank SafeGraph Inc. for permitting access to various SafeGraph datasets ('core places', 'weekly patterns'). Thank you also to Alessia Calafiore and Jacob Macdonald for their assistance and advice with the Open-Source Routing Machine (OSRM hereafter) and H3 spatial indexing system respectively.

4. Integrating the *Who, What* and *Where* of U.S. Retail Center Geographies.

The content of this chapter is published as a research paper in Annals of the American Association of Geographers:

Ballantyne, P., Singleton, A., Dolega, L., Macdonald, J. 2022. Integrating the Who, What and Where of American Retail Center Geographies. *Annals of the American Association of Geographers.*

DOI: https://doi.org/10.1080/24694452.2022.2098087.

Note: language in this chapter reflects the fact that this is published in an American journal (e.g., center instead of centre).

Chapter Overview

This chapter comprises the second of three empirical chapters in this thesis, and fulfils the second aim of this PhD thesis by generating a comprehensive overview of the geographies of the U.S. retail centre system. Using data from SafeGraph, retail centre boundaries (the *where*) are delineated using a new method based on H3, a higher resolution American retail centre typology is generated (the *what*), and retail centre catchments are extracted through calibration of a traditional Huff model with mobility data from SafeGraph to understand *who* is using these retail centres and where they come from. Furthermore, it is demonstrated that these three geographical aspects of retail centre geographies (the *who*, *what* and *where*) are better understood when considered together, through supporting the empirical insights gained, and present a new conceptual framework which better situates these empirical insights within the wider retail (centre) system, through conceptualisation of the interactions between different components, and their links to external pressures.

The contributions of this chapter are as follows. For the first time, a comprehensive model of the American retail centre system is provided, through capturing of *where* they are located, *what* characteristics they have and *who* uses them, through the use (and enhancement) of techniques in retail centre delineation, classification and probabilistic modelling. Secondly, the importance of integration in such analyses is demonstrated, through accounting for the intrinsic links between each of the three rerail centre geographies, demonstrating in particular the relationships between function and scale, and the continued applicability of the Huff model. Finally, the first conceptual framework of retail centre geographies is presented, which grounds such empirical insights in the context of the wider retail system, highlighting the apparent connections between retail centre geographies, and their relationship to external pressures.

Abstract

Retail is an important function at the core of urban areas, occupying a key role in determining their economic prosperity, desirability, and vibrancy. Efforts to understand the geographies of retail centers, the cores of retailing in urban areas, have a long academic tradition, often studied through either rich local case studies, or when geographically more expansive, are constrained by limited detail. New data in United States detailing the location and uses of retail creates a significant opportunity to develop a more complete and comprehensive overview of the national retail system, at a high spatial resolution. This research is rooted in a pragmatic effort to provide the first and most comprehensive model of U.S. retail center geographies, through development of an integrated, conceptual, and empirically grounded framework, using data from SafeGraph, to examine *where* they are located, *what* characteristics they have, and *who* uses them. The resulting geographies are of great interest, creating significant potential in the monitoring of the national retail system as it continues to evolve in response to wider structural challenges. Furthermore, by integrating these three geographies (*where*, *what*, and *who*), a conceptual framework is established that yields substantive insights about the relationships between each of them, and argues that understandings of U.S. retail center geographies are more comprehensive and useful when considering the *who*, *what*, and *where* together.

4.1. Introduction

The significance of retail to a nation's economic development has long been realized. As a vital economic contributor and supplier of employment (Helm et al., 2020), the retail system often acts as a barometer for wider economic trends (Berman and Evans, 2013). Furthermore, in the process of urban growth, the expansion of cities is in part related to their function as a consumer hub (Han et al., 2019), with the desirability of cities increasing with a greater diversity of consumption amenities and global service (Glaeser et al., 2001). Retail centers, the "main cores of retail activity in urban areas" (Dolega and Celińska-Janowicz, 2015, 9), are inextricably linked to the desirability of cities, as the primary sites of consumption within them. As a result, it is no surprise that policymakers increasingly view retail centers as integral to economic prosperity (McCann and Folta, 2008), despite mounting evidence that their economic value in cities and towns has declined in the past decade (Wrigley et al., 2015).

Despite this, research on the nationally extensive and comprehensive geographies of retail centers, in terms of their location, structure, and function, remains scarce (Sevtsuk, 2014). Such research has a wealth of benefits beyond simply understanding how they occupy space. Understanding the location patterns of retail (the *where*) provides policymakers and stakeholders with critical supporting information to maximize the economic output of the industry, through better urban economic policy (Larsson and Öner, 2014). Furthermore, discerning the provision and characteristics of retail centers (the *what*) has significance in understanding the composition, structure, and "vibrancy" of the retail system (DeLisle, 2005; Dolega et al., 2021), and in the creation of highly livable, mixed-use, and sustainable built environments, with a strong "sense of place" (Baker and Wood, 2010; Sevtsuk, 2014). When integrated with an understanding of *who* is using retail centers and *where* they come from, these geographical understandings can facilitate better evidence-led decisions about retail location, and the subsequent development of policy. Thus far, such insights into the geographies of U.S. retail centers have been relatively inconsistent (Baker and Wood, 2010), despite a common recognition that these localities and the social constructs they engender are vital.

In developing such understandings, coverage, and replicability are key. Up-to-date and expansive knowledge of the structure of retail provides tools to monitor the evolution of retailing (Joseph and Kuby, 2016), and therefore insights about different retail spaces and time periods (Guy, 1998). For example, reconstructing definitions and typologies can contribute substantive insights into the response of retail centers to challenges, such as the so-called 'retail apocalypse' (Helm et al., 2020). Combined with the unfolding long-term impacts of the COVID-19 pandemic, policy action is critical to protect the "brick-and-mortar" component of the U.S. retail system (Torres-Baron, 2018). For such policy to be feasible, however, there needs to be a comprehensive understanding of retail center geographies at the national level, which make it possible to both identify where policy is required and assess the effectiveness of such actions (Lloyd and Cheshire, 2017).

Literature on the geographies of retail centers is well established within the retail geography and urban studies communities, although often only considering a singular aspect of their geographies; for example, *where* they are located, *what* characteristics they have, and *who* uses them. Regarding *where*, recent examples have typically focused on development of analytical frameworks for delineation of the scale and extent (boundaries) of retail centers (Pavlis et al., 2018; Nong et al., 2019; Ballantyne et al., 2022a). In examining *what* characteristics they have, articles have placed an emphasis on trying to the understand the position of centers in the wider retail system, through development of typologies and classifications (Brown, 1992). There is a substantial legacy of academic enquiry seeking to understand patronage to stores, using empirical models to identify *who* uses them (Huff, 1964; Wilson, 1969; Fotheringham, 1983). Although understandings for retail centers are limited, the delineation of retail center catchments has recently become more feasible (Dolega et al., 2016).

Despite a wealth of literature considering these three retail center geographies in isolation, there has thus far generally been no consensus on how best to bring them together to provide a comprehensive overview of a national retail system (DeLisle, 2005). A key constraint is often the availability of data for the national extent; however, within the United States, new data are enabling such a vision to be realized, as demonstrated in the analytical framework proposed by Ballantyne et al. (2022a). In their article, the authors constructed a framework for extracting information about U.S. retail centers; the scale, extent, and characteristics of those in the Chicago Metropolitan Area. In this research, their framework is extended, with the aim of providing the first expansive, comprehensive, and fully integrated study and conceptual framework of the geographies of U.S. retail centers. In doing so, key technical contributions are made such as demonstrating the utility of non-hierarchical retail classifications, and the continued relevance of the Huff model, as well as significant substantive contributions when integrating the three geographies of U.S. retail centers, by demonstrating the connections between them.

Thus, this article has three key objectives: understanding the *who, what*, and *where* of U.S. retail centers. Following a review of existing literature on retail center geographies, the data sets used in this investigation are introduced, before outlining the approaches used and resulting geographies of the retail centers (the *where*), their characteristics (the *what*), and catchments (the *who*). In Section 4.7, the implications of this study are discussed, emphasizing the value of integration; understandings of the *who, what,* and *where* are strengthened when examined together.

4.2. Background

The Definitions and Origins of the Retail Center

The clustering or agglomeration of retail outlets in geographical space has been defined in various ways. A common definition used in the (U.S.) retail geography literature is shopping center or hub (Clapp et al., 2019; Rao, 2020). Such definitions, however, often only incorporate purpose-built retail developments, excluding unplanned clusters in decentralized locations, as well as other complementary functions (Guy, 1998). The term *retail center* arguably has greater saliency, in its consideration of all retail, its complementary functions (e.g., leisure), and an array of forms and functions. The significance of this definition was first recognized by the U.S. Census Bureau in 1966 (see Casparis, 1969), but definitions have since evolved, owing to advances in greater computational capacity and data availability (Pavlis et al., 2018).

A clear and conceptual definition was proposed by Dolega and Celińska-Janowicz (2015), "the main retail cores in urban areas," and although definitions between studies differ, there is one term that connects them all: agglomeration. Referring to a mass or collection of things, agglomeration remains one of, if not the key concept in understanding the locational behavior of retail activities (Brown, 1992; Sadahiro, 2000). Retail agglomerations occur because there are inherent benefits for retailers when locating in close geographical space (Zhou and Clapp, 2015), such as increased access to an existing stream of consumers (Teller and Reutterer, 2008), and significant cost advantages (Vom Hofe and Bhatta, 2007).

Thus, it is unsurprising that retail agglomerations are described as one of the most ubiquitous features of the commercial environment (Brown, 1992).

Retail center studies are also inextricably linked to CPT, which posits that distinctive hierarchical structures are in place between agglomerations, determined by the demand for which goods are purchased (Brown, 1992; Parr, 2017). Many nonhierarchical and irregular retail structures are emerging, however, (Dolega et al., 2021), and some previous assumptions about consumer behavior, made by CPT, are now considered less relevant (Brown, 1992). These theoretical considerations are important though, when demonstrating the need for posthierarchical understandings and awareness of the inapplicability of the spatial component of CPT (Borchert, 1998), crucial when undertaking analysis of *where* retail centers are located.

Where? A Short History of Retail Center Delineation

Responses seeking to determine *where* retail centers are, and the extent of geographical space that they occupy, can be traced back as far as 1928 (Woodbury, 1928). Traditionally, delineations of retail center space were based on field observation or mapping of features (Woodbury, 1928; Berry, 1963; Clark, 1967). Owing to advancements in the availability of expansive spatiotemporal data sets and analytical capabilities (Nong et al., 2019), however, contemporary examples have been able to take a more data-driven perspective, utilizing data aggregated to grid cells (Wang et al., 2014; Han et al., 2019) or the individual store locations themselves (Porta et al., 2009; Han et al., 2019) to explore the location distribution patterns of retail centers. Both Pavlis et al. (2018) and Ballantyne et al. (2022a) used spatial clustering of store locations to delineate robust retail center boundaries for the United Kingdom and Chicago Metropolitan Area. There were some notable limitations in both approaches, however, relating to the underlying data and computational capacity of the HDBSCAN algorithm, respectively. In their analytical framework, though, the authors demonstrated the effectiveness of the H3 spatial indexing system in refining center boundaries (Uber, 2018), which has also been used recently to define retail centers in the United Kingdom (Macdonald et al., 2022). Although it is clear that there is no global consensus on how best to delineate the *where* of retail centers, it is evident that computational cost and data reliability remain key constraints. If full replicability and coverage are to be achieved, the approach needs to be scalable and repeatable, something that an approach based on H3 can arguably provide.

What? Positioning Retail Centers in the Wider Retail System

Knowing *where* retail centers are located forms a necessary precursor to understanding *what* characteristics they have, through the development of typologies or classifications and hierarchies (Guy, 1998). These studies, which seek to understand different retail formats and the roles they occupy, are abundant historically and in a more contemporary sense, as retail systems are constantly evolving and transforming (Micu, 2019; Rao, 2020). Historically, classifications of these spaces were predominantly concerned with hierarchies and emphasizing the role of size in overall function (Sadahiro, 2000). Such approaches have retained saliency to the present day, for example in the international shopping center classification (ICSC, 2017).

In the past few decades, however, there have been some fundamental shifts in the way that people shop, reflecting the growing role of online channels and changing demands of consumers. As a result, new store formats have emerged (Joseph and Kuby, 2016) to support the demand for increasingly experiential retail and leisure (Dolega and Lord, 2020). Thus, for any measure of retail function to have contemporary relevance, it has to reflect such changes, still accounting for a greater multidimensionality of descriptive input measures to effectively differentiate between different functions (Guy, 1998; Rao, 2020). Such approaches are becoming increasingly feasible, owing to advancements in the way classification is conceptualized and measured in retail geography (e.g., Brown, 1992), but also to advancements in analytical capacity and data (Dolega et al., 2021), providing significant scope for advanced retail classifications based on sophisticated empirical analysis (DeLisle, 2005).

Who? Catchment Methodologies for Retail Centers

In understanding the geographies of retail centers, it is also vital to understand the demand: customers *who* are using them, and where are they are coming from, referred to as the *catchment*. Catchments can be defined as “the areal extent from which the main patrons of a store or retail centre will be found” (Dolega et al., 2016, p1), resulting from the spatial shopping behaviors of consumers, and influenced by a multitude of different factors. There are numerous ways to delineate catchments, but broadly they can be split into two categories. Deterministic methodologies, such as circular/fixed-buffer (Berry et al., 2016) and drive time or distance polygons (Rudavsky et al., 2009) have some advantages, but rarely capture the complexity of real-world consumer behavior and competition (Dolega et al., 2016). On the other hand, probabilistic methodologies include the entropy maximization, competing destination, spatial interaction, and Huff models (Wilson, 1969; Fotheringham, 1983; Newing et al., 2015; Dolega et al., 2016). The latter, based originally on the law of retail gravitation (Reilly, 1931), posited that consumer patronage could be modeled by considering the spatial distribution and attractiveness of locations, and accounting for relative competition between them (Huff, 1964).

A criticism of probabilistic models, however, is that they denote how retail patronage should occur, and as such do not always reflect more complex reality. Supported by advancements in customer-level data (Dramowicz, 2005), though, such models have become more sophisticated and flexible (Newing et al., 2015), calibrated with data on consumer behavior (e.g., loyalty cards) to ensure they generate more accurate representations. These calibration data sets are typically used to estimate where consumers come from (Waddington et al., 2018), enabling delineation of an approximate catchment (Davies et al., 2019). Despite these advancements, robust empirical catchment delineations for retail centers are sparse in the related literature (Pratt et al., 2014), likely relating to the additional considerations needed (Dolega et al., 2016). This presents a unique opportunity in the United States given the availability of highly accurate spatiotemporal consumer data (SafeGraph, Inc., 2020b), and advancements in parameter estimation methodologies (Wang et al., 2016; Liang et al., 2020).

Integrating the Who, What, and Where

Although there is a plethora of literature about these geographies in isolation, the relationships between all three in the wider retail system have received limited conceptualization, despite well-known connections between them. For instance, it is well understood that retail centers of differing scales exhibit significant differences in function, as historically modeled by CPT, and used as a prerequisite to most hierarchical classifications of consumption spaces (Berry, 1963; ICSC, 2017). Despite numerous studies using the outputs of the *where*—retail center boundaries—as inputs to generate the *what*—retail classifications—(Pavlis et al., 2018; Ballantyne et al., 2022a), through development of analytical frameworks, no direct theoretical consideration has been given to the connections between them.

Such connections are also apparent between the *what* (retail classifications) and the *who* (retail catchments). The relative supply of goods and services is one of the key determinants of retail patronage; centers deemed to provide a multipurpose or comparison-shopping experience typically draw consumers from a wider area than perhaps smaller town and district centers serving local communities (Guy, 1998; Dolega et al., 2016). Finally, there are also notable connections between the *who* and the *where*; the spatial location or distribution of retail centers determines the extent to which they compete with each other, which is also in part related to the function and hierarchical position of the retail center(s)—the *what*. Thus, it has been common to account for both function and scale when constructing catchments for a national system of retail centers; for example, by applying arbitrary values based on size and type of shopping center (ICSC, 2017), or by deploying Huff models for convenience and comparison centers separately (Dolega et al., 2016).

Thus, it is clear that these three retail center geographies (the *who, what,* and *where*) are intrinsically linked to each other, as presented within the conceptual framework (Figure 9), and can be better understood through integration, to provide a comprehensive overview of a national retail center system. Figure 9 conceptualizes the interactions between each of the three retail center geographies, as discussed earlier; for example, the relationship between scale and function, and the importance of function to modelling patronage, all of which are explored in the remaining sub-sections of this chapter. Figure 9 also highlights the pertinence of external pressures (e.g., online shopping), which occupy a significant role in the operation of a national retail system, and are thus closely linked to the three geographies. It is argued that better integrating the *who, what*, and *where* provides a more comprehensive overview of U.S. retail center geographies, and can thus help to better understand and effectively respond to external pressures on the retail system.

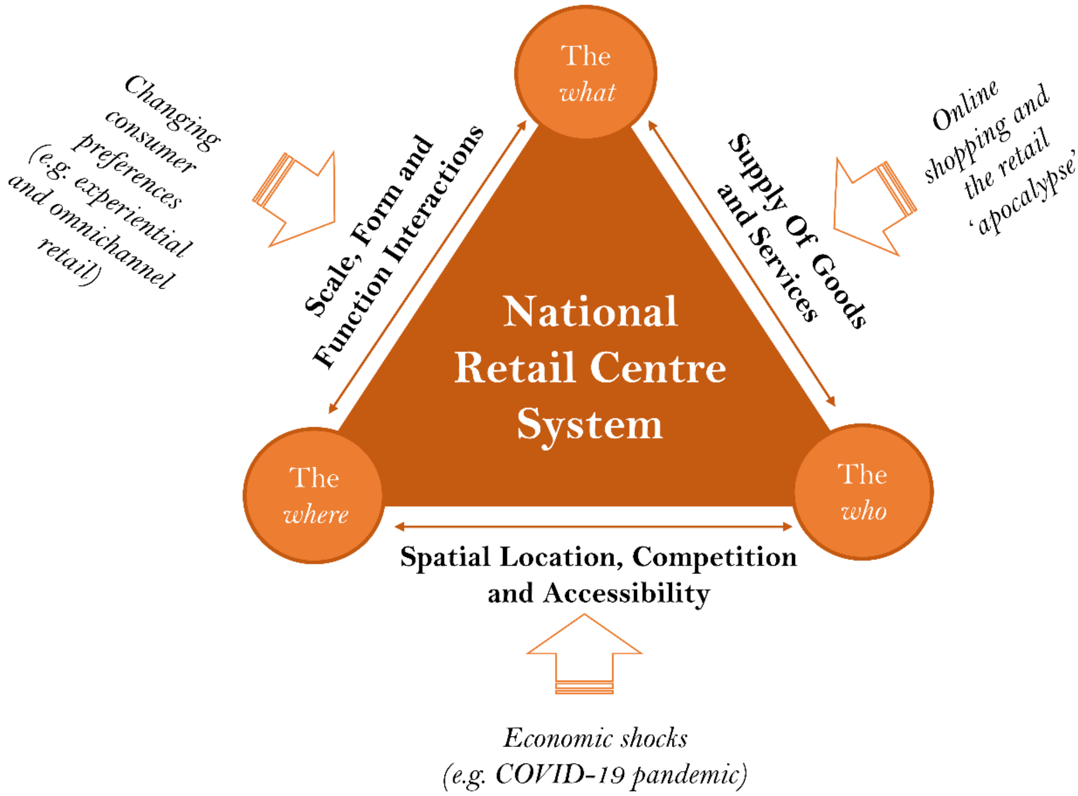

Figure 9. Conceptual framework of integrating the *who, what* and *where.*

4.3. Data

The primary database used in this research— "places"—was obtained from SafeGraph, a U.S. company that provides POI and mobility data for the United States, Canada, and the United Kingdom. This database contained information on places where people spend money, time, or both, made up of three primary data sets. The first data set, core places (SafeGraph Inc., 2020a), contained attribute information, such as the name, address, category, and coordinates. The geometry data set (SafeGraph, Inc., 2020c) contained detailed building footprints for each place, accounting for relationships between individual places and shared buildings. The final data set, patterns (SafeGraph, Inc., 2020b), contained traffic and demographic aggregations for the places, providing detailed information on how often people visit, how long they stay, and how far they travel. The SafeGraph places database was used as it is openly accessible, offers comprehensive coverage for the national extent of the United States, and as argued by Ballantyne et al. (2022a), is the best available source of data on retail location for the United States in terms of accuracy and comprehensibility. Thus, it is no surprise that a large body of pioneering research has utilized the SafeGraph places database to understand geographical phenomena (Liang et al., 2020; Wang et al., 2021; Yabe et al., 2021; Zhai et al., 2021; Huang et al., 2022).

For this study, the places that related to retail were extracted, as in the analytical framework first proposed by Ballantyne et al. (2022a), resulting in a data set of 3,476,542 retail places, a list from which, for the national extent of the United States, base information (core places), building footprints (geometry), and corresponding mobility data (patterns) could be extracted for the retail places. With the latter, the more recent SafeGraph weekly patterns data set was used, which is available and updated weekly. A summary of how the three data sets (core places, geometry, and weekly patterns) were used in each component of this article can be seen in Table 3. Additional ancillary data sets included the use of retail land-use polygons from OpenStreetMap (OSM hereafter) to capture the extent of larger retail developments (e.g., parking lots), and spatial data on waterbodies from the U.S. Geological Survey to account for the presence of major rivers in boundary delineation. In addition, data from the Environmental Protection Agency (EPA, 2021) and the U.S. Census Bureau (2021) were used to derive useful information about the surrounding urban morphology (e.g., road density) and neighborhood income of retail centers, respectively, when constructing the typology.

Table 3. Data set usage within each of the three components.

Component	Data source	Dataset	Usage
The *where*	SafeGraph	*'core places'* *'geometry'*	Aggregated to a grid of H3 polygons.
	OSM	Retail land-use polygons	
The *what*	SafeGraph	*'core places'*	Composition, diversity, size and function variables.
		'geometry'	Size and function variables.
		'weekly patterns'	Economic performance variables.
	Environment Protection Agency	Smart Location Database	Size and function variables.
	U.S. Census Bureau	American Community Survey	Economic health variables.
The *who*	SafeGraph	*'weekly patterns'*	Calculating observed patronage (origin census tract, total visits)

4.4. The *'Where'* of American Retail Center Geographies

Delineating the Scale and Extent of U.S. Retail Centers

The approach used to delineate retail centers used the H3 spatial indexing system, which as discussed earlier, holds significant potential for an accurate delineation that is both interoperable and computationally inexpensive. The approach, adapted from that used by Macdonald et al. (2022) and outlined in 10, takes as input the 3,476,543 retail places introduced earlier, their longitude/latitude coordinates from the core places data set, their building footprints from the geometry data set, and retail land-use polygons from OSM (Figure 10A). These input data sets were then aggregated to a grid of H3 addresses that contained retail features (Figure 10B); the chosen H3 resolution that the features were aggregated to was eleven, pertaining to hexagons with a diameter of 50 m, the optimal resolution for exploring connectivity between different retail places (Ballantyne et al. 2022a).

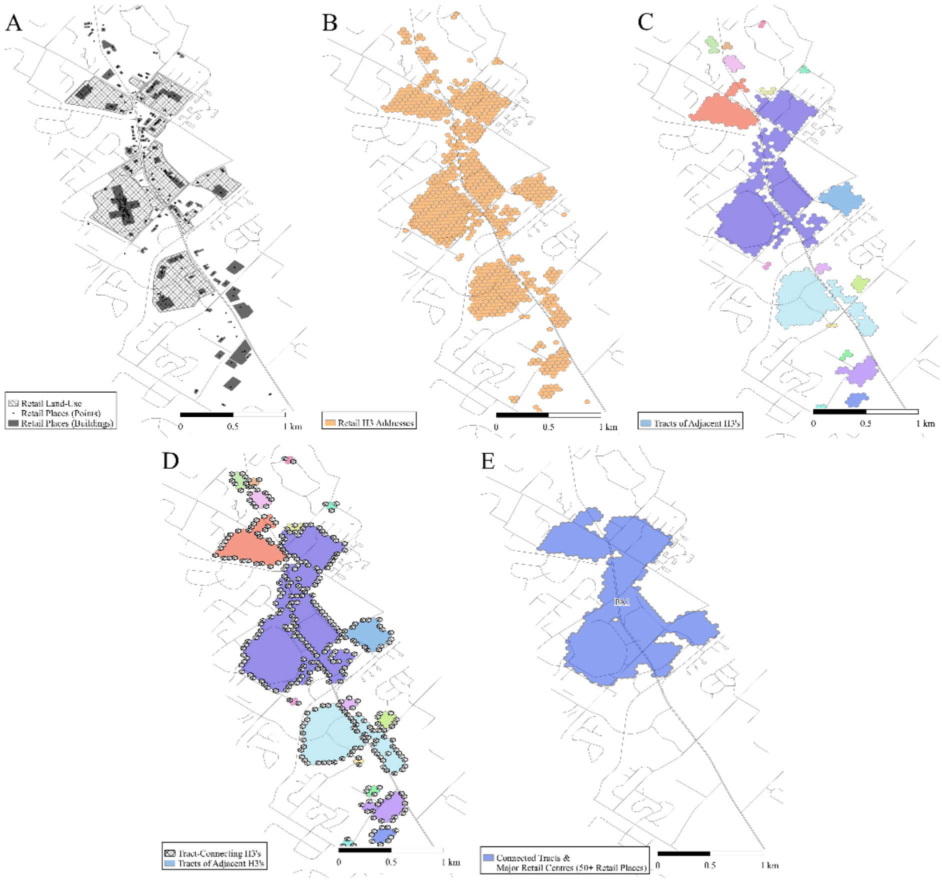

Figure 10. Approach to retail center delineation.

Once the various input data sets had been aggregated to H3, the retail center boundaries were delineated, beginning by assembling a series of contiguous tracts of hexagons based on direct adjacencies (Figure 10C), with only directly-neighboring hexagons in the same tract. Following this, a search was then performed to see if any of these initial tracts were near to others—within 50 meters—by extracting a series of tract-connecting hexagons (Figure 10D); H3 addresses in the respective neighboring “K-rings” of two hexagons, in and between different tracts. A detailed visualization of what a K-ring is and how this operation worked can be seen in Figure 11. These tract-connecting H3’s (Figure 10D) were then used to merge nearby tracts built on adjacencies, resulting in a set of connected tracts (Figure 10E). Additional components included the use of the National Hydrography Dataset to restrict merging of tracts across major rivers, a particularly pertinent issue in cities like Providence and Chicago. Roads and railway lines held similar potential, but were excluded due to a lack of detailed spatial data detailing the existence of overpasses and underground railway lines, which have a limited impact on boundaries. Computationally, the approach used h3jsr, an R package for performing spatial operations within H3, and the delineation of centers occurred at the state level to improve performance. Only major retail centers were extracted—those with more than fifty retail places—to remove the large numbers of small centers, which distort the retail center typology (Ballantyne et al., 2022), and reduce the computational efficiency of catchment extraction.

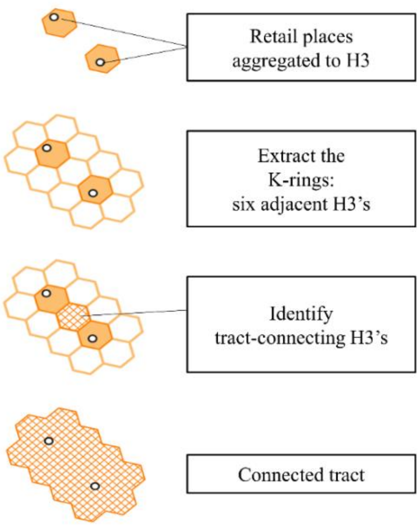

Figure 11. Use of K-rings to identify tract-connecting H3's and build connected tracts.

Where are U.S. Retail Centers?

The distribution of U.S. retail centers can be seen in Figure 12, comprising a total of 10,956 (major) retail centers. The majority were found almost exclusively in urbanized and heavily populated areas of country: 98 percent were within an official MSA, with the most populated MSAs containing the greatest numbers of centers (e.g., Los Angeles). The centers varied greatly in size, with the smallest containing fifty retail places, and the largest (Manhattan, New York) containing 27,907 retail places, with the median number being 85. These differences contribute to interesting debates about the continued role of scale in retail center geographies, something explored later. There were also interesting regional differences in the size and total number of centers; for example, despite containing the largest center (Manhattan, New York), the Northeast region comprised the smallest number of centers, with the greatest number of centers found in the South, which is realistic considering 40 percent of the population live there (U.S. Census Bureau 2021). It is important to note, however, that by excluding the smaller retail centers, this distribution is heavily skewed toward the most urbanized retail centers, as in Pavlis et al., (2018). By excluding these centers, it is likely that this study is not capturing their geographies in rural areas of the United States, a limitation that could not be avoided.

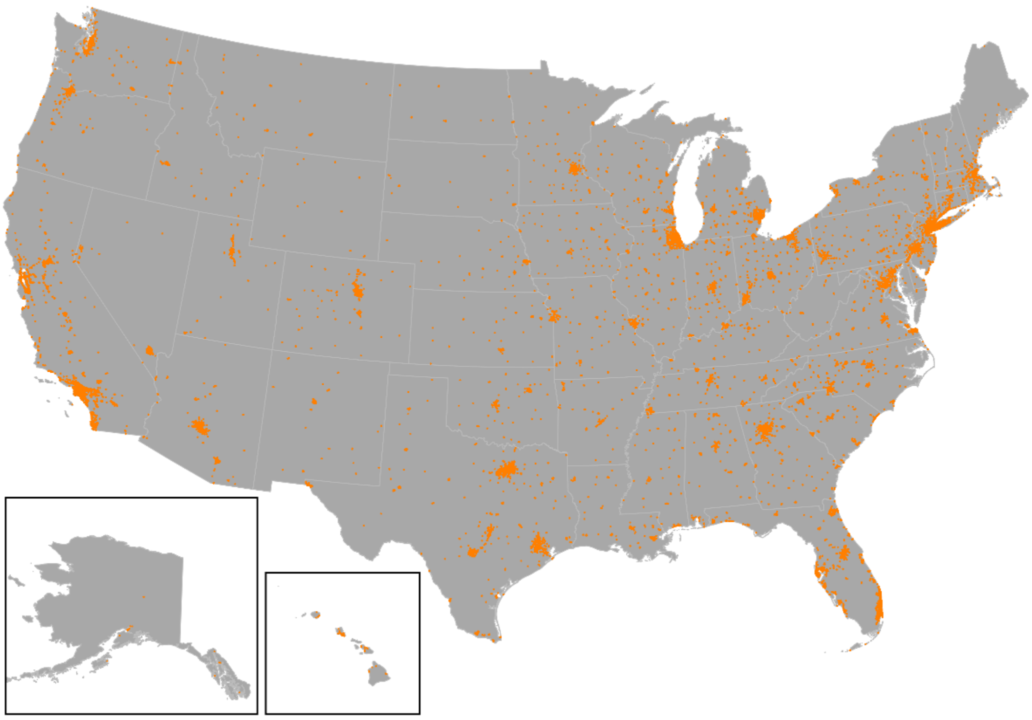

Figure 12. Distribution of U.S. Retail Centers (map not to scale).

The distributions of retail centers for some contrasting urban and retail environments are shown in Figure 13, where a broad range of forms are apparent including large, sprawling centers, such as those in Chicago and San Francisco (Figures 13A, 13D), and highly linear ones such as Downtown Boulder (Figure 13B). The existence of these different forms has long been recognized, with Berry (1963) distinguishing them as “ribbons” and “centers” in Chicago, whereas in the United Kingdom, Pavlis et al. (2018) likened them to “chain” and “compact” centers. The large sprawling retail centers (Figures 13A, 13D) occurred in many other U.S. cities such as Seattle and Washington, similar to those seen in UK cities like Liverpool or Manchester (Pavlis et al., 2018), and similarly, the linear centers (Figure 13B) are not all that different from UK high streets. At this point, it is worth noting that a number of methodological advancements enabled accurate delineation of the retail center boundaries. The use of building footprints over point data and water bodies has arguably better captured the spatial distribution of retail, without privileging its relationship to streets or major rivers. In addition, the use of land-use polygons enabled delineation of major retail developments and shopping centers typically enclosed by large parking lots, as in Figure 13C.

4.5.The *‘What’* of American Retail Center Geographies

Developing a Multidimensional U.S. Retail Center Typology

To account for functional differences between the retail centers, a multidimensional retail center typology was constructed. A series of variables (Table 4) were extracted for each of the four retail classification domains—composition, diversity, size and function, and economic health—as in the analytical framework first proposed by Dolega et al. (2021). The variables were mostly derived from the core places data set, with others from either the SafeGraph weekly patterns data set (SafeGraph Inc., 2020b), or other ancillary data sources (EPA, 2021; U.S. Census Bureau, 2021). The framework is extended, however, to ensure greater applicability to the United States, through capturing of a greater number and variety of measures, specific to the U.S. retail context, a problem identified by Ballantyne et al. (2022a). The final list of variables and their descriptive statistics can be seen in Table 4, and a detailed overview of how they were constructed can be found in Appendix II.

Figure 13. The *where* of U.S. retail centers in four contrasting urban and retail environments (maps not to scale).

Table 4. Variables used in construction of the retail center typology.

Domain	Variable	Descriptives	
		Median	Standard deviation
Composition	propClothingandFootwear (%)	6.00	10.37
	propDIYandHousehold (%)	3.21	2.91
	propElectrical (%)	*1.27*	*1.32*
	propRecreational (%)	3.99	3.19
	propChemist (%)	1.96	2.01
	propCTNandGasoline (%)	3.28	2.45
	propFoodandDrink (%)	4.51	2.89
	propGeneralMerchandise (%)	*0.96*	*1.47*
	propBars (%)	1.13	2.81
	propRestaurant (%)	17.80	8.32
	propFastFood (%)	7.27	4.48
	propEntertainment (%)	1.72	3.49
	propFitness (%)	1.89	2.49
	propConsumerServices (%)	26.33	10.53
	propHouseholdServices (%)	4.62	6.23
	propBusinessServices (%)	3.47	4.25
Diversity	propIndependent (%)	43.59	19.52
	propSmallMultiple (%)	1.37	1.93
	propNationalChain (%)	*23.21*	*16.81*
	propPopularComparisonBrands (%)	1.23	6.47
	propPopularConvenienceBrands (%)	5.66	3.87
	propPopularLeisureBrands (%)	0.39	1.12
	nationalRetailDiversity (%)	*13.68*	*4.56*
	nationalServiceDiversity (%)	*13.51*	*5.31*
	localRetailDiversity (%)	37.84	12.98
	localServiceDiversity (%)	37.04	14.02
Size & Function	nUnits	85.00	371.93
	nBuildings	*82.00*	*393.93*
	area (km^2)	0.42	0.84
	roeckScore	0.37	0.21
	medianDistance (km)	8.30	175.56
	retailDensity	0.21	0.15
	residentialDensity	2.35	4.79
	retailemploymentDensity	*0.28*	*0.45*
	roadDensity	16.43	6.51
	propAnchor (%)	0.28	0.94
	propPremiumBrand (%)	0.00	1.92
	propDiscount (%)	1.08	0.43

Economic Health	totalVisits	3289.50	11867.31
	totalPopulation	*5864.50*	*15400.11*
	medianUnemployed (%)	*2.88*	*2.13*
	medianIncome ($)	58503.75	30047.21
	retailService	*1.99*	*21.44*
	nCompeting	40.00	77.42

To ensure that the input variables were parsimonious, correlation and sensitivity analysis were used to remove highly colinear variables, and those with little effect on the classification shape. Seven variables exhibited high collinearity and were removed, a key step when assembling classifications with unsupervised machine-learning techniques (Singleton et al., 2020). PCA was then used to identify those variables making limited contributions to the classification, which resulted in the removal of a further three variables. The final set of thirty-seven variables were range standardized (0–1), and the classification was performed using the PAM algorithm. PAM requires specification of the number of clusters (*k*); here elbow plots were used in conjunction with average silhouette scores to determine *k* in each iteration of PAM, which can be seen for the groups in Figure 14. The classification was performed twice to extract a two-tier classification, comprising a set of four retail center groups and a series of nested types.

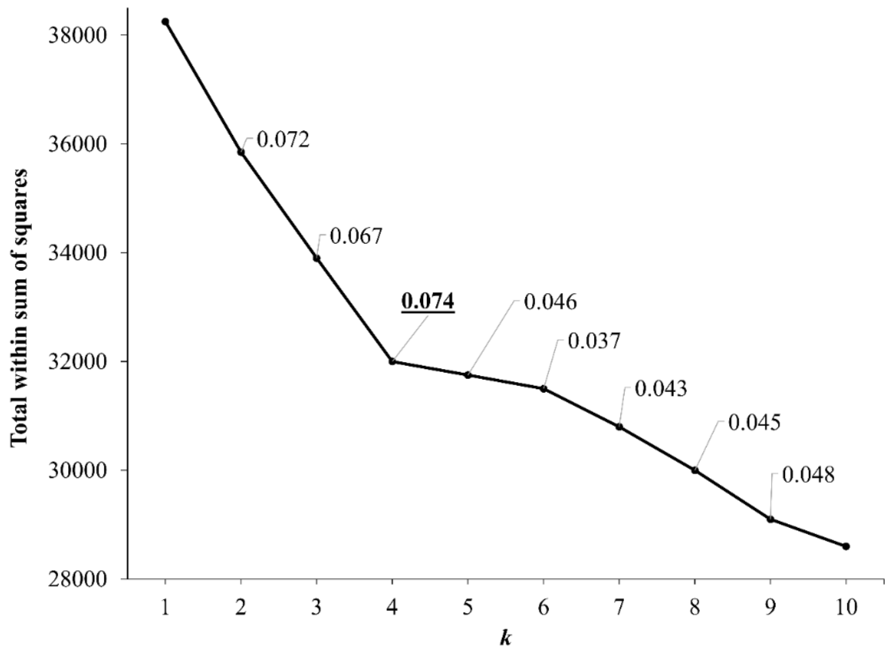

Figure 14. Determining the optimal *k* value for retail center groups. Value in bold represents optimal *k* value.

What Are the Characteristics of U.S. Retail Centers?

The characteristics of the retail centers, seen in Figure 15 and Table 5, were determined by considering variability in the median values of input variables between each retail center group and type. The spatial distribution of groups (Figure 16) was interesting; the abundance and uneven distribution of groups 2 (Small city, town, and primary neighborhood) and 4 ("Everyday" convenience and service) was particularly noticeable, forming clusters in and around smaller cities and other urban areas. In contrast, there were distinct concentrations of group 1 centers (Major urban centers and established shopping destinations) in the major cities such as Los Angeles and New York, with other isolated occurrences in smaller, state capital cities. The number of centers in each group varied substantially from 4,929 (group 4) to 1,316 centers (group 3, Leading comparison destinations). This difference does, however, seem plausible given the frequency at which everyday goods (e.g., groceries) are purchased, when compared to other types of retail goods (e.g., home furnishings). Furthermore, when considering the total number of major cities versus smaller cities and towns in the United States, the greater number of centers in group 2 compared with group 1 also seems plausible.

1. Major urban centres & established shopping destinations (n = 1,606)

- Large, non-specialist, highly-diverse.
- Abundance of independents, bars, restaurants and services.
- Heavily-urbanised locations (high road, residential and building densities).
- Lower dominance of discount stores and popular brands.

2. Small city, town and primary neighbourhood centres (n = 2,745)

- Abundance of independents, bars, restaurants and services.
- Highly linear morphology.
- Smaller, less diverse and located in less urbanised locations than group 1 centres.
- Lower dominance of popular brands and anchor stores.

3. Comparison and multipurpose shopping destinations (n = 1,316)

- High popularity and diversity.
- Abundance of popular comparison brands (e.g. fashion) and anchor stores.
- Good mix of discount stores and premium brands.
- Typically located away from heavily urbanised areas (low road/residential density).

4. 'Everyday' convenience and service retail centres (n = 4,929)

- Abundance of food and drink, service and convenience retailers.
- Visited frequently (popular).
- Anchored by a general merchandiser (e.g. Walmart), numerous discounters.
- Located away from dense urban areas, much smaller, less abundance of popular comparison brands.

Figure 15. Pen portraits for the retail center groups.

In terms of composition and function (Figure 15), interesting differences were apparent between the groups. For instance, both the first and second group of centers comprised a retail offering typical of urban centers (e.g., abundance of bars, restaurants). They were differentiated, however, by size, diversity and urban morphology, resulting in a clear distinction between major city centers (e.g., Chicago, Illinois) and smaller urban centers (e.g., Aurora, Illinois). Furthermore, the distinction between Groups 3 and 4 was of great interest; despite polarized retail offerings, discount and anchor (general merchandise) retailers occupied a key role in both. When examining the nested types (Table 5), many interesting functional and compositional differences were also identified. The existence of some of these types within the groups was logical, for instance the splitting of premium outlets and leading fashion destinations (3.2) from off-price, nonspecialist comparison destinations (3.4). Furthermore, the differentiation of primary and secondary metropolitan centers (1.1, 1.2) and the identification of food and drink destinations (1.3). As concluded earlier, though, it is important to note that the typology presented here likely excludes some additional retail center functions, particularly those exclusive or common in rural areas.

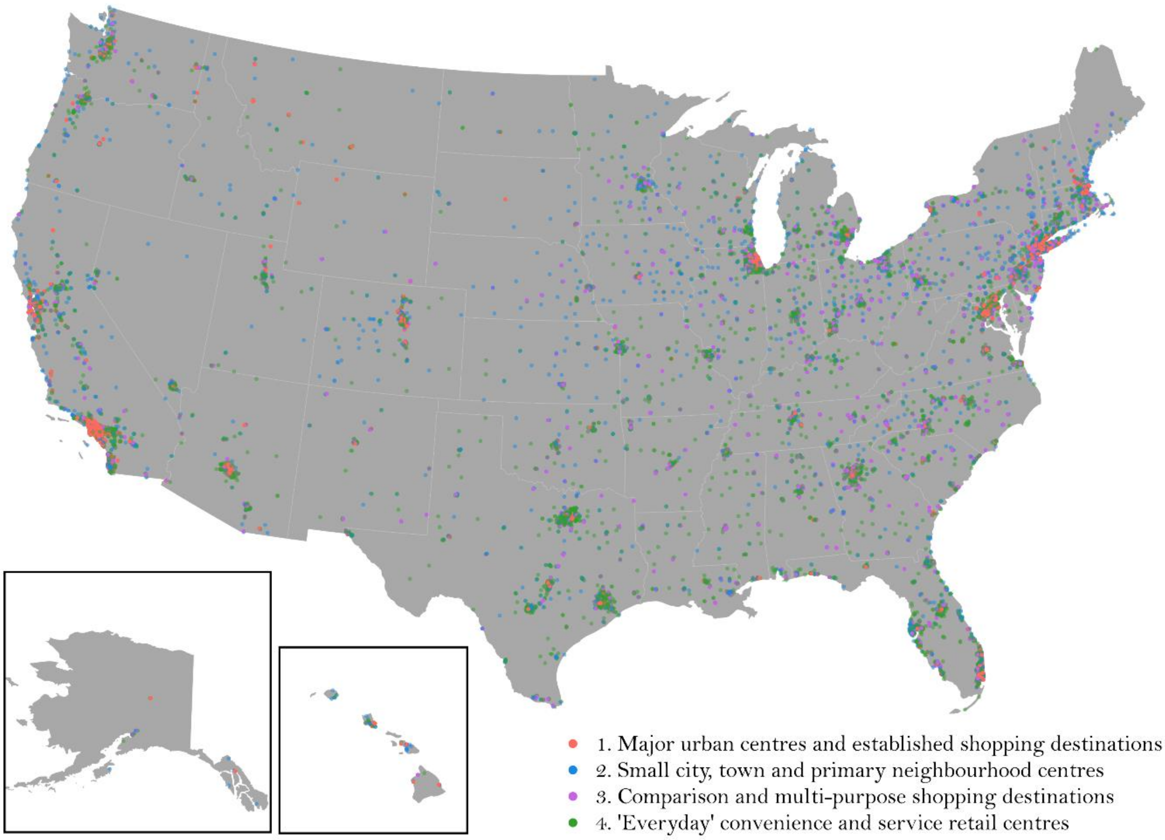

Figure 16. Distribution of U.S. retail center groups (map not to scale).

Here a multidimensional retail center typology for the United States has been constructed, which highlights the apparent structural and functional interdependencies between centers. Despite a non-hierarchical approach, the relationship between function and scale must not be overlooked, as highlighted in the conceptual framework (Figure 9). Dividing up the retail center groups into four size categories (Table 6) yielded interesting insights about the mapping of functions across various scales, and the continued applicability of CPT. For instance, with the smaller and medium-sized retail centers, a greater diversity of functions was evident, particularly an abundance of localized ones, as suggested by Guy (1998). When looking to the large and very largest retail centers, it is apparent that there are fewer centers in these categories with a much lower overall diversity. These centers were typically found in the largest U.S. cities where CPT is arguably of much less contemporary relevance, as polycentricity and existence of large spatial structures creates significant market fragmentation (Dolega et al., 2016). Thus, the vast differences in retail center functions and scales across the United States counter efforts to conceptualize the retail system through CPT, as paralleled in the United Kingdom (Dolega et al., 2016), but must not be overstated as the smallest centers have been deliberately excluded from this analysis.

Table 5. Characteristics of the U.S. retail center types.

Type	Key characteristics	Examples
1.1. Metropolitan and primary urban centres	Largest in terms of area and number of retail places, most diverse and most popular.	Manhattan (NY), San Francisco (CA), The Loop, Chicago (IL), New Orleans (LA).
1.2 Secondary metropolitan centres and iconic shopping districts	Smaller and less popular than 1.1, but much more diverse and popular than types 1.3 and 1.4.	Newark (NJ), Anchorage (AK), Portland (ME), Pittsburgh (PA), Berkeley (CA).
1.3 Inner city food and drink destinations	Smaller than other centres in group one, with a higher than average abundance of independents, restaurants and bars.	Venice, Los Angeles (CA), Pilsen & Logan Square, Chicago (IL), Rittenhouse Square, Philadelphia (PA).
1.4 Multipurpose peripheral shopping areas	Located in periphery of major cities, large and highly diverse, higher than average numbers of household services.	Hackensack (NJ), Watertown (MA), West End, Atlanta (GA), Ocotillo Plaza, Las Vegas (NV).
2.1. Small city, primary neighbourhood cores and secondary food and drink destinations	More diverse, greater proportion of leisure-based retail (e.g. bars), dominance of independent retailers.	Elgin (IL), Idaho Falls (ID), Fayetteville (AR), Fort Myers (FL), Bloomington (IN), Wailea (HI).
2.2. Neighbourhood specialist service centres	Less diverse, greater proportion of consumer services, less independents and bars/restaraunts.	Thompson Lane Center, Nashville (TN), Capital Square, Raleigh (NC), Gloversville (NY).
3.1 Large, popular and multipurpose destinations	Large, lots of weekly visits, with high number of comparison brands, services, anchors and discounters.	Shoppers World, Framingham (MA), Milford Crossing, Milford (CT), Ingram Park Village, San Antonio (TX).
3.2. Premium outlets and leading fashion destinations	Dominance of clothing and footwear, with an abundance of premium and the most popular comparison brands.	Orlando vineland premium outlets (FL), King of Prussia mall (PA), Waikele premium outlets (HI).

3.3. Secondary fashion shopping destinations	Abundance of comparison brands and clothing/footwear, department store anchored, absence of premium brands.	Mondawin Mall, Baltimore (MD), The Promenade, Bolingbrook (IL), The Summit, Reno (NV).
3.4. Off-price, non-specialist comparison destinations	Lots of discounters and anchor stores, non-specialist retail offering (e.g. comparison, convenience, fast food).	Assembly square, Somerville (MA), Fairlane green, Detroit (MI), Southern hills plaza, Oklahoma City (OK).
4.1. Affluent 'everyday' centres	High income neighbourhoods, lots of consumer services, convenience goods and fitness facilities.	Glen Gate, Morton Grove (IL), Port Jefferson Shopping Plaza (NY), Malibu Village, Malibu (CA).
4.2 Large and diverse 'everyday' centres	Large, highly diverse, abundance of convenience retail/consumer services, relatively affluent neighbourhoods.	Kent Station, Seattle (WA), Lakewood City Commons, Denver (CO), Williamson Square, Franklin (TN).
4.3. Popular discount convenience centres	High weekly visits, abundance of discount retailers, specialism in convenience retail.	Sunshine Center, Panthersville (GA), Essex Junction Center, Essex (VT), Southern Blvd, Rio Rancho (NM).
4.4. Secondary discount convenience centres	Similar to 4.3 centres, slightly lower abundance of discounters, greater prevalence of consumer services.	Aspen Square, Laramie (WY), Michigan Avenue, Detroit (MI), Wampanoag Plaza, Providence (RI).

Table 6. The observed relationship between U.S. retail center scale and function.

Retail centre group	**Retail centre size (Fisher-Jenks)**			
	Small (*nUnits* = 50 > 250)	Medium-sized (*nUnits = 250 > 1200*)	Large (*nUnits* = 1200 > 3200)	Very large (*nUnits* >= 3200)
	Percentage			
1. Major urban centres and established shopping destinations	10.49	76.70	100.00	100.00
2. Small city, town and primary neighbourhood centres	27.67	2.18	0.00	0.00
3. Comparison and multi-purpose shopping destinations	12.43	13.10	0.00	0.00
4. 'Everyday' convenience and service retail centres	49.41	8.01	0.00	0.00
Total retail centres (=100%)	9865	687	39	5

4.6. The '*Who'* of American Retail Center Geographies

Calibrating the Huff Model to Estimate Retail Center Catchments

Unpacking *who* uses retail (centers) and where they come from is a long-standing theme in retail geography. It is vital for better understanding the interplay of supply and demand, in particular, comprehending the main drivers of demand and access (Waddington et al., 2018). In this application, the approach of Dolega et al. (2016) was modified, where a bespoke Huff model was developed to delineate catchments for retail centers in the United Kingdom. The Huff model, as outlined earlier and specified below in Equation 1, posits that consumer patronage can be modeled by considering the attractiveness (A_j) and spatial location or distance of retail locations (D_{ij}), with α and β calibration parameters used to ensure the model accurately represents reality.

In this research the Huff model is applied to the United States, calibrating it with SafeGraph's weekly patterns data set (SafeGraph, Inc., 2020b), and accounting for functional differences between retail centers. The data set, as described previously, contains aggregated visit counts for each retail place at the census tract level enabling identification of *who* uses these places and (approximately) where they come from. Arguably, the weekly patterns data set could be used to directly demarcate catchments, but this is problematic, as catchments remain inherently dependent on the data itself, raising issues of representativeness and universal applicability. On the other hand, the use of the patterns data and recent advances in calibration (Liang et al., 2020) could yield substantive insights about the drivers of patronage in the contemporary retail system (e.g., distance, attractiveness), through direct calibration of the model α and β parameters, demonstrating the (in)applicability of the Huff model.

In this approach, the Huff model was used to provide a probabilistic breakdown of the likelihood consumers from census tract *i* would visit retail center *j*, as specified in Equation 1. To measure the attractiveness of retail centers (A_j), an aggregate score was derived from a series of variables (Equation 1) deemed important for retail center attractiveness (Dolega et al., 2016; Gong et al. 2021). Euclidean distances (D_{ij}) were calculated between the census tracts and centers, as shortest network distances to all tracts were not computationally feasible.

$$P_{ij} = \frac{A_j^{\alpha} D_{ij}^{-\beta}}{\sum_{j=1}^{n} A_j^{\alpha} D_{ij}^{-\beta}}$$

Equation 1. Huff model specification.

Where P_{ij} is the probability that consumers located in census tract *i* would visit retail center *j*; A_j is the measure of attractiveness for retail center *j*, based on size, total visits, diversity of retail offer and presence of popular comparison brands; D_{ij} is the shortest Euclidean distance from census tract *i* to retail center *j*; α is the attractiveness parameter, determined through comparison with observed patronage; and β is the distance decay parameter, determined through comparison with observed patronage.

The final step in fitting the basic Huff model was to calibrate the model parameters (α, β). In Dolega et al. (2016), the authors determined α and β using related literature and survey observations. In this study, however, recent advancements in (data-driven) parameter calibration were utilized (Wang et al., 2016; Liang et al., 2020), comparing a series of Huff models to observed patronage behaviors, to empirically derive α and β values. The SafeGraph patterns data set was used to compute observed patronage probabilities, and then compare these to a series of Huff models with different α and β values containing predicted Huff probabilities for each census tract, as shown in Figure 17. Correlation testing was performed between the observed and predicted probabilities, as in Wang et al. (2016), to identify the calibrated parameters, as shown in Table 7. Once determined, these calibrated parameters (α, β) were used to delineate catchments for the retail centers, by extracting predicted probabilities above 50 percent and 25 percent as the primary and secondary catchments, respectively.

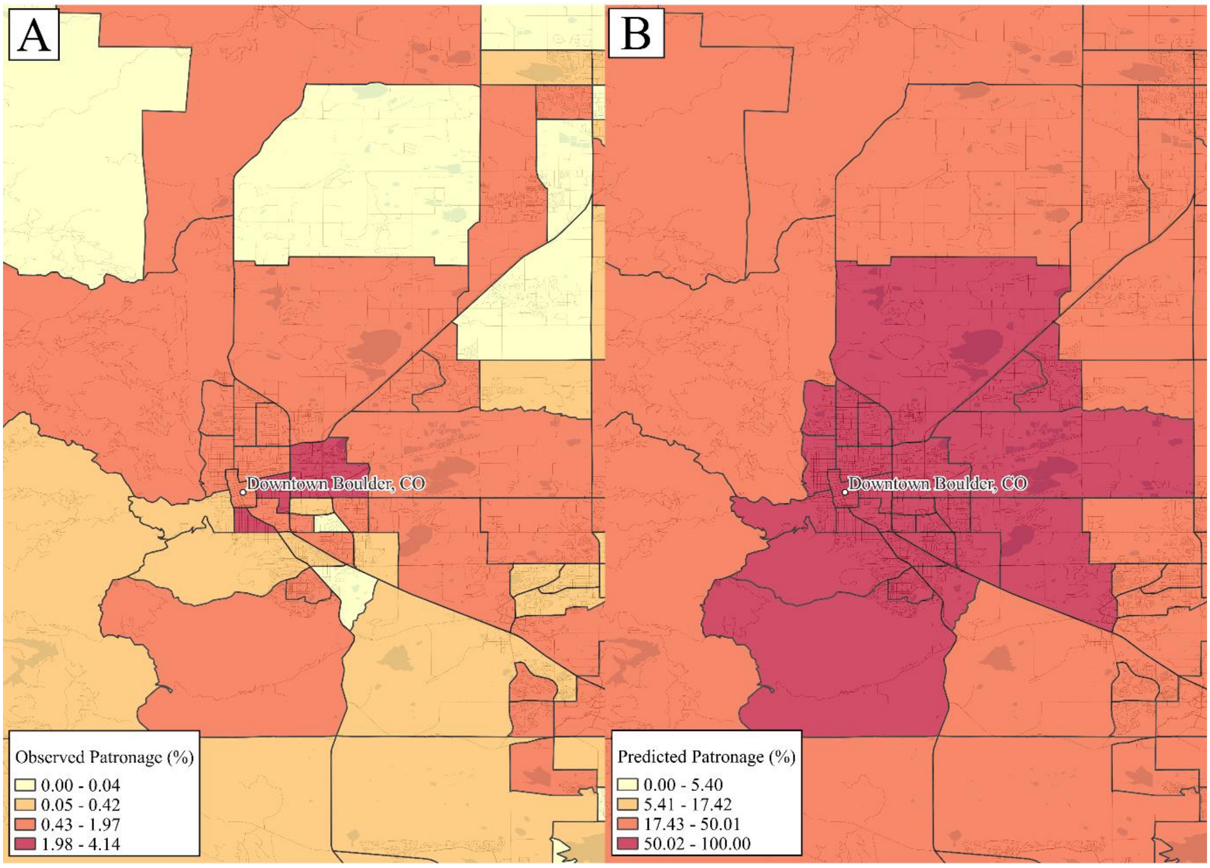

Figure 17. Comparing the (A) observed and (B) predicted patronage probabilities for Downtown Boulder (maps not to scale).

Table 7. Correlation testing of observed probabilities against predicted probabilities from Huff models with different alpha and beta values.

	β				
	Pearson's R				
α	0.1	0.5	1.0	2.0	5.0
0.1	0.309*	0.539*	0.657*	**0.665***	0.588*
0.5	0.270*	0.537*	0.654*	0.664*	0.588*
1.0	0.196*	0.529*	0.649*	0.663*	0.588*
2.0	0.109*	0.496*	0.632*	0.658*	0.587*
5.0	0.039*	0.339*	0.543*	0.626*	0.580*

As suggested by Dolega et al. (2016), a number of modifications and considerations were made to ensure suitability of this method to retail centers and maximize computational efficiency. The calibration of model parameters was performed for the West region of the United States, to minimize computational cost, with this region selected due to its diverse urban structure, comprising 30 percent of all centers. Second, this method was applied separately for each distinct type of retail center, ensuring only those with directly competing offerings were treated as equal on the catchment surface, as it is problematic to design a hierarchical catchment system (as in Dolega et al. 2016), based on a nonhierarchical typology. This also links to the conceptual framework (Figure 9), where it is illustrated that the *who* and *what* are intrinsically connected when considering that the supply of goods and services (the *what*) has a significant role in determining patronage (the *who*). Thus, a separate constrained Huff model was calibrated and used to extract catchments for centers in each type.

Who Uses U.S. Retail Centers?

The calibrated model parameters varied substantially between the retail center types (Table 8), offering useful insights as to the role of attractiveness and distance in determining patronage to U.S. centers with different functions. For example, much larger β values seen for group 4 centers were interesting, with these centers providing an “everyday” retail offering, a retail function highly sensitive to distance (Dennis et al., 2002). In addition, α and β were equal for the large, popular and multipurpose destinations (3.1). Given this type comprised many of the established U.S. shopping locations, they could be more likely to fit the conceptual basis on which the Huff model is grounded. In general, the attractiveness parameter (α) was of less significance in ensuring that the Huff models accurately represented reality, as in the United

Kingdom (Dolega et al., 2016), with β often exceeding α. What remains clear, however, is that a one-size-fits-all approach does not work for a national set of retail agglomerations, as in the United Kingdom (Dolega et al., 2016). To demonstrate this further, Figure 18 compares the catchment of a retail center that has (Figure 18A) and has not (Figure 18B) accounted for retail center type. With the former, a catchment that more accurately reflects the observed patronage behaviors was delineated, suggesting the role of function remains much more significant in determining patronage, over scale.

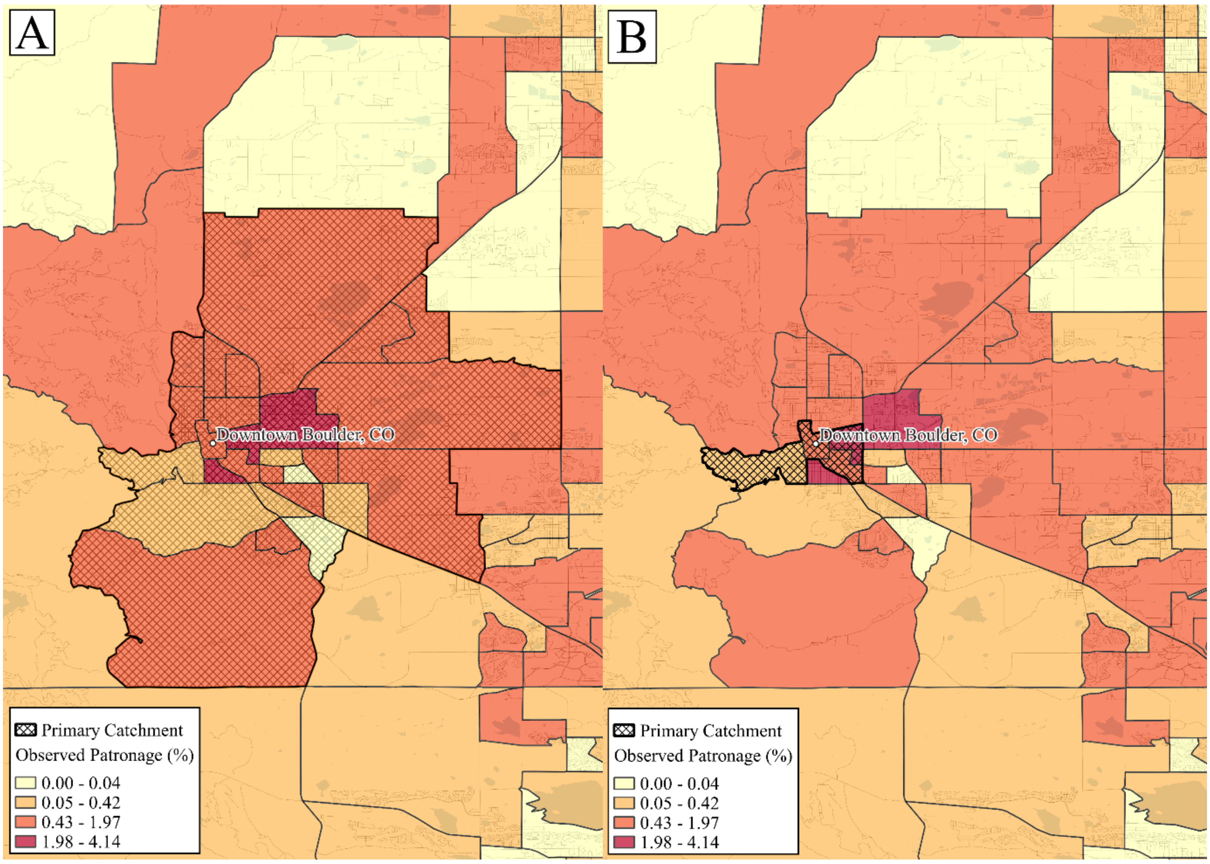

Figure 18. Primary catchment for Downtown Boulder, where the Huff model (A) has and (B) has not accounted for retail center type.

Table 8. Calibrated Huff model parameters for each retail center type.

Group	Type	α	β
1. Major urban centres and established shopping destinations	1.1. Metropolitan and primary urban centres	0.1	1.0
	1.2 Secondary metropolitan centres and iconic shopping districts	0.1	2.0
	1.3 Inner city food and drink destinations	0.1	1.0
	1.4 Multipurpose peripheral shopping areas	0.1	1.0
2. Small city, town and primary neighbourhood centres	2.1. Small city, primary neighbourhood cores and secondary food and drink destinations	0.5	1.0
	2.2. Neighbourhood specialist service centres	0.1	2.0
3. Comparison and multi-purpose shopping destinations	3.1 Large, popular and multipurpose destinations	1.0	1.0
	3.2. Premium outlets and leading fashion destinations	0.1	0.5
	3.3. Secondary fashion shopping destinations	0.5	1.0
	3.4. Off-price, non-specialist comparison destinations	0.5	1.0
4. 'Everyday' convenience and service retail centres	4.1. Affluent 'everyday' centres	0.1	2.0
	4.2 Large and diverse 'everyday' centres	0.1	1.0
	4.3. Popular discount convenience centres	0.1	1.0
	4.4. Secondary discount convenience centres	0.1	1.0

Primary and secondary catchments for Downtown Boulder and the Seattle City retail centers can be seen in Figure 19. For Seattle, its catchment was very typical of others in Type 1.1, typically very large, owing to a lack of directly competing centers nearby. In some of the more polycentric cities like Los Angeles and Chicago, however, the overall catchment sizes of these centers were much smaller, due to increased competition with other urban centers. Similarly, the Downtown Boulder catchments were large, with the nearest main competitor located in Denver, approximately thirty miles away. It is, however, likely that the Downtown Boulder retail center competes in some way with the nearby large 28th Street retail center, thus, an approach to catchment modeling at the group level might have handled competition more effectively. On experimentation, though, this failed to demarcate the naturally “higher order” Type 1.1 and 1.2 centers from others in the group, prompting further investigation into how to better measure the attractiveness of retail centers, and reduce the need for catchment overestimation.

Thus, although the role of competition has not been fully captured in the model, this study has demonstrated and calibrated a non-hierarchical Huff model that accounts for the function of centers entirely. Given that these catchments are calibrated against a large mobility data set, they arguably provide an accurate way of estimating patronage to retail centers for the national extent, despite overestimation. The underlying observed patronage behaviors highlight this (Figure 19), where it is evident that the majority of tracts containing relatively high levels of patronage were contained within the catchments of centers. One exception to this was the catchments for Type 3.2 centers, premium outlets and leading fashion destinations, which were the least accurate. Premium outlets, however, typically exhibit patronage behaviors that are distinctly different from all other comparison destinations (Guy, 1998); thus, it is likely that the Huff model in its current form is not sufficient to account for this difference in function.

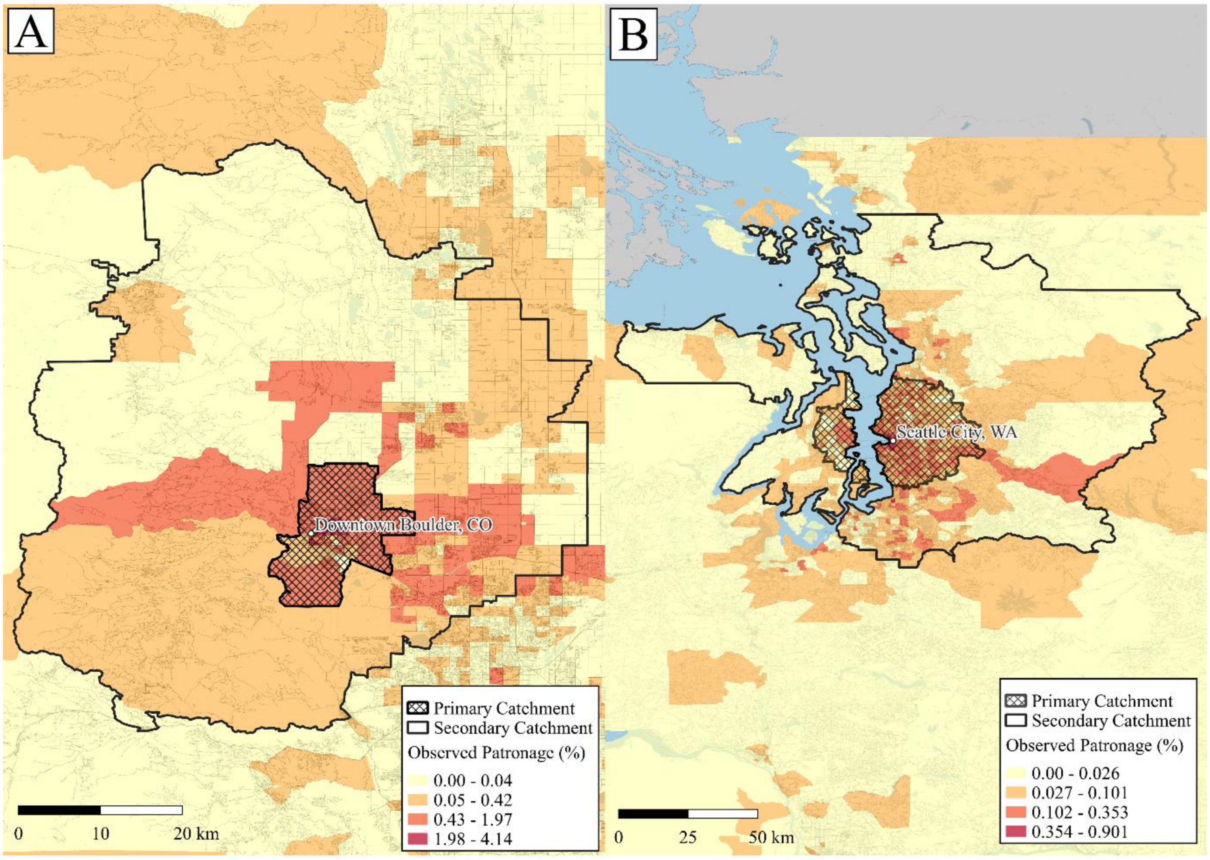

Figure 19. Primary and secondary catchments for the (A) Downtown Boulder and (B) Seattle City retail centers.

An unexpected, yet interesting aspect of the catchments was in their average sizes. Unsurprisingly, some of the traditionally higher order centers like Type 1.1, 1.2, and 3.1 were the largest in area, with the smallest catchments seen for centers providing an “everyday” or convenience-oriented offering (2.2, 4.1, 4.2). These findings have major implications for research like ours, which uses data-driven approaches to understand geographical phenomena. With retail centers there are always implicit hierarchies, but without local, expert knowledge, it becomes difficult to account for and build replicable approaches that integrate such knowledge. In this study, early steps have been made to demonstrate the potential for unpacking the hierarchical from the non-hierarchical, using a Huff model calibrated on a large mobility data set to identify higher and lower order patronage behaviors, demonstrating further the connection between the *who* and the *what*.

4.7. Discussion and Conclusions

The physical, brick-and-mortar component of the U.S. retail sector is under threat. Retail centers, the primary sites of (physical) consumption (Dolega and Celińska-Janowicz, 2015), are under increased pressure during the so-called retail apocalypse, and as longer term, structural impacts continue to disrupt the retail landscape—not least given the recent shifts due to the COVID-19 pandemic. This research is rooted in a pragmatic effort to better understand the geographies of U.S. retail centers as a response to these issues, through extension of the analytical frameworks first proposed in Dolega et al. (2021) and Ballantyne et al. (2022a), to provide a new conceptual framework that yields substantive insights about the U.S. national retail center system. Using data from SafeGraph and cutting-edge techniques in retail center delineation, classification, and probabilistic modeling, three geographical aspects of U.S. retail centers are explored*: where* they are located, *what* characteristics they have, and *who* uses them. In developing such understandings, though, it is argued that these three geographical aspects are intrinsically linked, and as such can be better understood when examined together, through provision of a conceptual framework to ground such understandings.

For instance, the efficacy of including ancillary data sets to derive better retail center boundary delineations was demonstrated, resulting in a higher resolution and more representative retail center typology than was obtained in Ballantyne et al. (2022a). Furthermore, it was demonstrated that the Huff model can be enhanced to better account for implicit differences in patronage between centers, by integrating information about retail function, supplied by typologies. Finally, throughout this article, the apparent connections between scale and function are highlighted, illustrating that retail center functions can span multiple scales, and that function remains a greater determinant of patronage over scale. Thus, empirical inquiries into the geographies of U.S. retail centers need to be better integrated, considering the *where, what,* and *who* together, to derive more substantive and useful insights, as opposed to considering them in isolation or duality, as in much of the related literature.

Fundamentally, these apparent links between the spatial distribution, typologies, and catchments of U.S. retail centers contribute heavily to theoretical and conceptual underpinnings in retail geography, particularly in comprehending the role of function and scale in the retail system, and the utility of fit for purpose nonhierarchical classifications, where critical details and niches about retail environments have now been captured. Furthermore, the continued applicability of the Huff model in retail (center) geographies was evidenced, through successful calibration of the model using a large mobility data set from SafeGraph, Inc. (2020b), also shedding significant light on the changing role of function, attractiveness, and distance in conceptualizing patronage.

These conceptual contributions echo many of the findings of other studies in Annals of the American Association of Geographers, notably Scharadin et al. (2022) and Shannon (2016), who investigated the geographies of food environments. Both studies noted the role of food offering and function(s) in determining the patronage behaviors of (retail) food environments, calling for more holistic and multidimensional understandings of them, and highlighting the apparent utility of using observed mobilities to better understand and model these patronage behaviors; enabling definitions of neighborhoods based on individual movements across space and time (Root, 2012). Thus, although these studies did not explicitly integrate the three geographies of food environments—*where, what,* and *who*—they also demonstrated the intrinsic links between them. This research, however, provides significant scope for future studies into food environment geographies, both in demonstrating the utility of mobility data, to derive catchments or trade areas for a larger number of food environments beyond those based on individual observations (as in Shannon, 2016), and more broadly, providing a systematic framework through which to measure the national geographies of food environments.

As a resource, these U.S. retail center geographies also offer significant potential in helping to identify how and where effective responses are needed, to protect the physical component of the U.S. retail system, and the social and economic value that they represent (Lloyd and Cheshire, 2017). Given the role of retail centers in affecting the livability and desirability of cities (Sevtsuk, 2014), this new knowledge can be used to support development of legislation and the design of cities (Baker and Wood, 2010). Placing an emphasis on the overall sense of place and the quality of the retail offer can result in significant enhancements to the livability and economic success of these areas (Glaeser et al., 2001; Sevtsuk, 2014). Furthermore, given the continued role of the pandemic in our daily lives, increasing volume of online sales and the expanding network of literature on the retail apocalypse, developing metrics for the centers presented here, such as those seen in Singleton et al. (2016) and Comber et al. (2020), could provide an assessment of the pertinence of these issues across the entire U.S. retail system.

These structural challenges are not unique, as they are pertinent in other international settings, which can also benefit from such geographical understandings, and the development of effective, data-driven policy action in response. For such outcomes to be feasible, however, the empirical measurement of retail center geographies has to be replicable (Dolega et al., 2021), offering repeatability in existing contexts, and applicability to different ones. Through the development of a fully replicable workflow, utilizing open-source tools and methodologies, and creation of a comprehensive GitHub repository (Ballantyne, 2022), the geographies of U.S. retail centers can be updated at regular intervals, enabling insights about their evolution to be gained (Joseph and Kuby, 2016). Furthermore, in conjunction with the increased availability of globally available retail location data (Safegraph Inc., 2021b), and a new conceptual framework providing a comprehensive overview of the national retail (center) system (Figure 9), the workflow presented here can be modified and extended to derive impactful understandings of retail center geographies in other international settings, as suggested by Ballantyne et al. (2022a).

This research does not claim to provide the definitive set of U.S. retail center geographies, as it is inherently limited by a lack of engagement and validation involving stakeholders, local experts, and qualitative understandings. Furthermore, a significant limitation is in exclusion of the smallest and likely most ruralized retail centers. These are likely to exhibit a significantly different distribution when considering *where* they are located, and as a result, all retail center functions have likely not been captured and described. This also has notable implications when considering the relationships between retail center scale and function, as discussed earlier. Further studies into U.S. retail center geographies at the national extent should seek to explore these ideas further, seeking to understand what additional knowledge can be generated about the geographies of the U.S. retail system when incorporating these localities. Limitations aside, it is argued that this study has been able to provide an empirically grounded and conceptual framework through which to better understand the geographies of U.S. retail centers. Throughout this article, new knowledge has been generated about *where* they are located, *what* characteristics they have and *who* uses them, and more importantly emphasized the importance of integration, utilizing the conceptual framework presented here, to yield the most compelling and useful geographical insights about these phenomena.

Supplementary Material

To see the accompanying supplementary materials for this chapter (and published paper), please see Appendix II.

Acknowledgements

The authors would like to thank SafeGraph, Inc., for permitting access to various SafeGraph data sets through their academic program. We would also like to thank Professor Ling Bian (editor) and the two anonymous reviewers for their kind and thoughtful feedback on this article at various stages.

5. Using unstable data from mobile phone applications to examine recent trajectories of retail centre recovery.

The content of this paper has recently been accepted for publishing as a regular research paper at Urban Informatics:

Ballantyne, P., Singleton, A., Dolega, L. Using unstable data from mobile phone applications to examine recent trajectories of retail centre recovery. *Urban Informatics* (accepted).

Chapter Overview

This chapter comprises the final of three empirical chapters in this thesis, and fulfils the third aim of this PhD thesis by exploring how retail centres have recovered following the COVID-19 pandemic from summer 2021 to 2022. In particular, mobility data from Geolytix is used to identify significant heterogeneities in recovery between retail centres in England, Scotland and Wales, with noticeable differences when explored at the regional level or between different types of retail centres (i.e., *functions*). However, following identification that significant variation exists with regions and *functions,* modelling revealed that the *structural* characteristics of retail centres appeared to be more closely associated with recovery, notably the resilience of retail centres to online shopping, catchment deprivation and the composition of leisure and service retailers. This chapter also yields significant insights about the utility of data derived from mobile phone applications, particularly about the key temporal limitations of the Geolytix mobility dataset, identifying ways in which to utilise such unstable data to derive meaningful insights.

The key contributions of this chapter are as follows. Firstly, the potential of the Geolytix mobility dataset for spatio-temporal analysis is evaluated, as it has not yet been used in published research. In particular prominent temporal limitations are identified which restrict its use in examination of temporal trends, but can be used as a comparative tool when treated as snapshots, assuming representativeness and stability between compared areas are formally considered. Secondly, substantive insights about the drivers of retail centre recovery through exploration and modelling are presented, identifying significant heterogeneities, with particularly strong associations between the *structural* characteristics of retail centres and their likelihood to recover. Finally, an empirical basis upon which to monitor the recovery of retail centres is provided, in particular calling for greater multidimensionality in how retail centre performance in conceptualised, and utilising some of the tools presented to yield such insights with other unstable sources of data.

Abstract

The COVID-19 pandemic has changed the ways in which we shop, with significant impacts on retail and consumption spaces. Yet, empirical evidence of these impacts, specifically at the national level, or focusing on latter periods of the pandemic remain notably absent. Using a large spatio-temporal mobility dataset, which exhibits significant temporal instability, the recovery of retail centres from summer 2021 to 2022 is explored, considering in particular how these responses are determined by the *functional* and *structural* characteristics of retail centres and their regional geography. These findings provide important empirical evidence of the multidimensionality of retail centre recovery, highlighting in particular the importance of composition, e-resilience and catchment deprivation in determining such trajectories, and identifying key retail centre *functions* and regions that appear to be recovering faster than others. In addition, a use case for mobility data that exhibits temporal stability is presented, highlighting the benefits of viewing mobility data as a series of snapshots rather than a complete time series. Such data, when controlling for temporal (in)stability, can provide a useful way to monitor the economic performance of retail centres over time, providing evidence that can inform policy decisions, and support interventions to both acute and longer-term issues in the retail sector.

5.1. Introduction

The COVID-19 pandemic has caused significant damage across societies and economies around the world (Duong et al., 2022). As a result of policy actions imposed at various stages to mitigate the spread of the disease, the pandemic has severely disrupted daily activities, and has, and continues to change those ways in which we shop (Sit et al., 2022). This has had notable consequences for physical spaces of consumption such as high streets and retail centres, which have struggled for many years prior to the pandemic (Dolega and Lord, 2020). Within the UK, and in advance of COVID-19, vacancy rates were at an all-time high since the 2008 economic crisis (Wrigley et al., 2015), and footfall was significantly down (HSTF, 2021), in part due to the increasing popularity of online shopping and out-of-town shopping centres (Enoch et al., 2022). However, there is now a growing evidence base that the pandemic has accelerated these trends, often being likened to a 'pandemic retail apocalypse' or 'catalyst for change' (Frago, 2021).

Despite a wealth of literature exploring the short and medium-term impacts of public health restrictions on the retail sector (Baker et al., 2020; Nicola et al., 2020; Bonaccorsi et al. 2020), there has thus far been limited efforts to directly quantify these responses for retail centres, accounting for spatial heterogeneities at the regional level, and their *functional* and *structural* characteristics. The focus of this paper is therefore on British retail centres – "the primary sites of consumption in urban areas" (Dolega and Celińska-Janowicz, 2015, p.9), and their recovery from the initial shock of the COVID-19 pandemic. Although some examples for retail centres have emerged in cities (Frago, 2021; Ballantyne et al., 2022a), these studies have emphasised the consequences of public health restrictions on retail centre activity, with much less written about the more recent 'phases', such as the Omicron variant. The latter is of great interest, as the Omicron subvariant re-infected many of those who were already vaccinated or had previously tested positive (Chowdhury et al., 2022; Grabowski et al., 2022), but saw no further public health restrictions, only recommendations.

In addition, existing studies have utilised various forms of data to assess the economic performance of consumption spaces, such as vacancies (Frago, 2021; Dolega and Lord, 2020) and footfall (Philp et al., 2022; Ntounis et al., 2020). The utility of mobility data in answering such questions was first identified in Trasberg and Cheshire (2021), providing significant scope for the use of similar data, such as the Geolytix aggregated in-app location dataset (CDRC, 2021a), to unpack how such responses have manifested in later phases of the

COVID-19 pandemic once the key limitations of such data have been addressed. Thus, within this context, the utility of a mobility dataset for exploring spatio-temporal trends of retail centre recovery is evaluated, demonstrating in particular *how* retail centre definitions and new forms of data can be used as geographic data tools, to better understand the response of the wider retail sector to the pandemic. As such, three research aims are proposed:

i) Consider the utility of the Geolytix mobility dataset for spatio-temporal analysis of retail centre recovery.
ii) Explore the extent to which these recovery trajectories relate to the overall *function*, and regional geography of retail centres.
iii) Quantify the role of the *structural* characteristics of retail centres, in addition to *function* and regional geography, in determining such recovery trajectories.

5.2. Background

The British retail (centre) landscape

The British retail landscape has undergone a large transformation. Driven in part by the rising popularity of 'E-commerce' (ONS, 2022), the expansion of out-of-town developments and economic 'shocks' like the 2008 recession (Dolega and Lord, 2020), we have seen a significant decline of traditional high streets and retail centres (Wrigley et al., 2015). As a result, vacancy rates are at an all-time high, with increasing unemployment and concentration of retail away from high streets (Jones and Livingstone, 2018; Parker et al., 2017). The COVID-19 pandemic represents another challenge, significantly reducing footfall in many consumption spaces, following implementation of mobility restrictions to contain the spread of the virus (Enoch et al., 2022). Whilst these factors are well acknowledged as being some of the primary drivers of 'brick-and-mortar' retail decline, research suggests that these impacts are spatially heterogenous, with retail (centre) vulnerability and decline being highly variable, driven by multiple factors related to the *structur*al and *functional* attributes and catchment characteristics of the centres (Dolega and Lord, 2020; Singleton et al., 2016). What remains clear however is that the decline of retail centres is a multidimensional issue, which becomes increasingly convoluted when studied at different spatial scales, highlighting the complexity and diversity of the problem (Parker et al., 2017).

Measuring retail centre performance

Although complex to capture, there is widespread consensus that data-driven empirical measures of performance hold great value for policy and planning of the future of cities and retail (Enoch et al., 2022; Philp et al., 2022). There are however no uniform indicators for measuring retail centre performance (Dolega and Lord, 2020), owing to the complexity of such a measure, and the influences of internal and external factors (Philp et al., 2022), as well as demand and supply (Jones et al., 2022). Total spend would be of greatest utility, but is difficult to obtain or estimate given the decentralised nature of retail(er) organisation. As such, there are numerous proxy measures that have been used, such as vacancy rates (Dolega and Lord, 2020; Jones et al., 2022), footfall (Philp et al., 2022; Ntounis et al., 2020) or attractiveness and retail mix (Dolega et al., 2016; Jones et al., 2022). However, such measures are subject to limitations, such as overly privileging certain geographic areas or having limited temporal resolution.

The increasing availability of new forms of data, creates novel opportunities for the monitoring of human mobility (Calafiore et al., 2022), and derivation of proxy performance measures for different places and spaces (Ballantyne et al., 2021), through which to understand urban problems. A large body of research is emerging that uses mobility data obtained from mobile phone applications to investigate human behaviour during the COVID-19 pandemic, notably changes in human mobility and internal migration (Kang et al., 2020), and the compliance of social distancing measures (Oliver et al., 2020). In addition, such data has been used to monitor the performance of consumption spaces and the wider retail sector during the pandemic (Trasberg and Cheshire, 2021; Ballantyne et al., 2022a; Ballantyne, 2021).

However, mobility data is not without limitations. Location data from smartphones face similar challenges to other consumer datasets in that they are often unrepresentative of particular social groups (e.g., generational biases), or of particular areas due to differences in access to mobile devices/internet (Trasberg and Cheshire, 2021; Parsons, 2020). In addition, such data often faces significant temporal limitations, depending on the sample of devices and applications used to collect it, and their representativeness of the general population (Gibbs et al., 2021). Typically, the panel of unique devices will vary over time, which must be accounted for when using such data to conduct any spatio-temporal analysis (Trasberg and Cheshire, 2021). Thus, such mobility data is subject to its own limitations, and uncertainty in the generalisation of any results generated remains a significant challenge (Shi et al., 2022; Gibbs et al., 2021). However, there is still more to be unpacked about how such datasets can be used to monitor the economic performance of retail centres, particularly their post-pandemic recovery trajectories, and how these link to their overall *functiona*l, regional and *structural* characteristics.

Retail centre performance and recovery

Observations about the short-term responses of retail centres to the pandemic, and the different restrictions and rules, form an essential basis to informing the preparedness of these locales in the future (Enoch et al., 2022). Much literature has focused on the consequences of restrictions during the earliest stages of the pandemic here in the UK. For example, Ntounis et al. (2020) and HSTF (2021) documented significant decreases in national footfall, whilst others identified notable disparities between different retailers (Baker et al., 2020; Nicola et al., 2020). Recently however, studies have emerged that examined these trends between different spaces of consumption, such as Enoch et al. (2022), who identified significant differences in footfall declines between UK town centres. Of great interest is how these disparities of impact and recovery relate to the characteristics of the retail centres, in particular their *functional* role (i.e., hierarchical positioning) and *structural* characteristics (e.g., vacancy rates). With *function,* studies have identified significant differences in responses between smaller, local centres and larger towns and cities (HSTF, 2021; Enoch et al., 2022; Ballantyne et al., 2022a; Frago, 2021), relating these trends to the role of commuting, goods, or scale of demand in determining such responses. With *structure,* research has identified significantly different responses depending on the composition, vacancy rate, resilience to online shopping (e-resilience hereafter), diversity of retail offer and catchment deprivation of different retail centres (Enoch et al., 2022; HSTF, 2021; Dolega and Lord, 2020). Furthermore, related research has argued that such responses will exhibit significant spatial heterogeneities (Dolega and Lord, 2020), thus the importance of geographical location (e.g., regional geography) cannot be overlooked.

However, all of the above examples have examined the responses of retail centres to the earliest ‘phases’ of the COVID-19 pandemic, with much less written about more recent ‘phases’, such as that seen over the past year, where the Omicron subvariant has been of great significance. In the UK, the pandemic has been characterised by different sets of restrictions during different time periods, in response to different variants of the original virus. However, following Omicron, the government unveiled a much less stringent set of restrictions – “Plan B”, comprising mandatory face masks and vaccine passports (Prime Minister’s Office, 2021), with these being lifted in January. Thus, the likely supply side impacts on consumption spaces were greatly curtailed in comparison to restrictions seen earlier in the pandemic, theoretically enabling recovery to begin at the start of 2022, though this may have differed between different countries (e.g., Scotland, Wales). Thus, the overarching objective of this paper is to examine how retail centres responded beyond national lockdowns, how these relationships map into recovery (or decline) trajectories across different *functional*, regional and *structural* characteristics, and the utility of mobility data for capturing such trends.

5.3. Data and Analysis

Geolytix 'aggregated in-app location dataset'

The primary dataset used in this research; Geolytix *'aggregated in-app location dataset'*, was obtained from the Consumer Data Research Centre (CDRC, 2021a). The dataset contains aggregated activity counts derived from in-app mobile phone applications across Great Britain, which are aggregated into a hexagonal geometry (H3), providing a count of the total number of distinct devices within each 50m hexagonal cell. The data provides hourly, daily and weekly counts, spanning a 365-day period from August 2021 to July 2022, with the best spatial coverage occurring in towns, cities and other urbanised areas. It is important to note however that it is impossible to identify the specific sources of data used to construct it (i.e., apps), as that information is commercially sensitive (CDRC, 2021a). For the purposes of this research in examining the response of retail centre activity, and to minimise disclosure risk, the mobility dataset was appended to the latest iteration of the CDRC retail centre boundaries (Macdonald et al., 2022); the nested H3 cells within each centre boundary were derived and joined with the corresponding mobility data, keeping only data within the centre boundary, before calculating the total number of devices within each retail centre at the weekly scale, as a proxy measure for retail centre activity, to smooth variation at the daily level.

However, the temporal stability of the Geolytix mobility data remains a significant limitation, as is often the case with other similar mobility datasets (Trasberg and Cheshire, 2021). As demonstrated below in Figure 20A, the number of unique devices in the Geolytix mobility dataset does not remain consistent throughout the entire study period, falling from around 170,000 in August 2021 to 65,000 by July 2022, for reasons which are unavailable as users rather than data providers or creators. Thus, it is no surprise that a decreasing number of devices in the sample over time results in decreasing average device numbers within retail centres, as below in Figure 20B. This raises significant questions about the suitability of the Geolytix mobility dataset for analysis of trends over time, as any temporal trends are likely to be heavily affected by the decreasing number of devices in the sample. However, upon consultation with Geolytix, it was suggested that this decrease of devices does not compromise the representativeness of different geographical areas (i.e., regions) and types of retail centre (i.e., *functions*), when examining temporal trends at short time-periods such as weeks. Evidence of this can be seen below in Section 5.4 where a representativeness and stability analysis is undertaken of the Geolytix data, before concluding that its stability between different regions and *functions* enables robust comparisons between retail centres in Section 5.4.

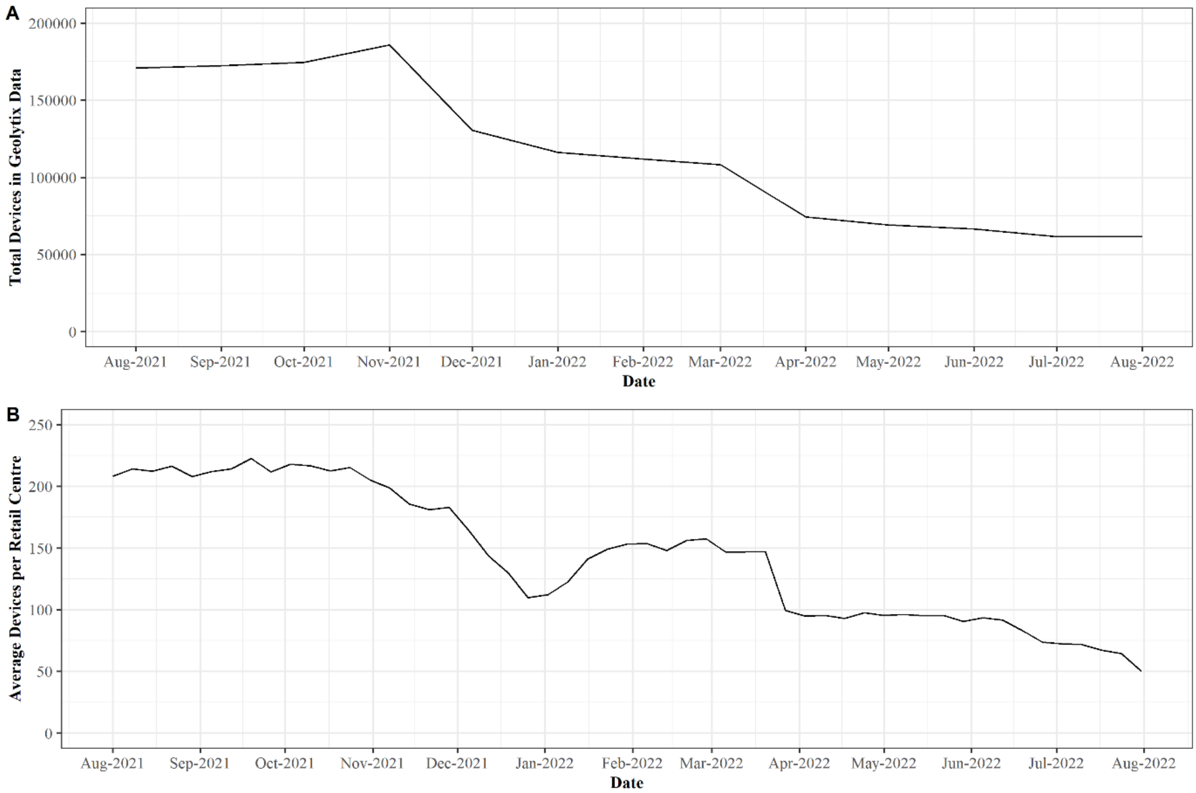

Figure 20. Changes in the number of devices in the Geolytix aggregated in-app location dataset throughout the study period, where A) demonstrates the falling number of devices in the sample and B) highlights its implications on the average number of devices within retail centres at the weekly scale.

Supporting datasets

To investigate the role of retail centre *function* and *structure* in determining the response of retail centre activity during the study period, the safeguarded CDRC 'retail centre indicators' data product was utilised (CDRC, 2021b), which provides summary indicators for the retail centres. Specifically, to characterise the *function* of the centres, the retail centre hierarchy

(Classification), as described below in Table 9, was utilised. To characterise the *structure* of the centres, the remaining indicators were utilised, listed below in Table 10, comprising information about the composition, diversity, catchment deprivation and e-resilience of the retail centres. As the retail centre indicators are available only for a subset of the 6,423 retail centres in Great Britain, only those retail centres for which both *functional* and *structural* indicators were available were used in this investigation. This results in exclusion of the large number of small centres across Great Britain and some additional larger centres (retail parks, shopping centres), for which indicators are not available and/or are significantly different in *function* and *structure* (CDRC, 2021b; Jones et al., 2022). These exclusions were not desirable; however, this step was unavoidable as disclosure risk means these indicators are not available for these very small retail centres. The result was a set of 1,068 study retail centres across the UK, comprising weekly data on retail centre activity (i.e., total devices) and the accompanying *functional* and *structural* indicators for the retail centres.

Table 9. Retail centre hierarchy (Classification), describing *functional* differences between the retail centres, obtained from CDRC (2021b).

Classification	**Examples**	**N**
Regional Centre	London, Birmingham City, Liverpool City, Manchester City, Glasgow City.	14
Major Town Centre	Carlisle, Warrington, Luton, Bournemouth, Swansea.	82
Town Centre	Grimsby, Welwyn Garden City, Clapham Junction, Torquay, Tenby.	270
District Centre	Ellesmere Port, Camden Town, Chesham, Greenside.	228
Market Town	Berkhamstead, West Kirby, Bakewell, Kenilworth, Billericay.	112
Local Centre	Newport Pagnell, Frodsham, Oadby, Egham.	378

Analytical approach

The economic performance of consumption spaces is a product of numerous forces of change, making it a highly complex problem to understand (Parker et al., 2017), and as highlighted thus far, existing research shows that the *functional* role, *structural* composition and regional geography are all linked to the overall performance of retail centres both in the short and longer term. Thus, in Section 5.4, following formal validation that the number of devices remained stable at the weekly scale and between different regions and *functions* as suggested by the data provider, this study explores how retail centres with differing *functions* and in different regions (see Figure 21) have responded during this study period, examining changes to activity within them. In particular, once specific regional biases created by the mobility data have been controlled for, changes to retail centre activity as share change between different *functions* and regions are examined, as it would not be appropriate to visualise change in total or average devices over time, as these trends would be subject to underlying limitations of the data (Section 5.3).

Table 10. Retail centre *structural* indicators, obtained from CDRC (2021b).

Variables	**Description**
propChain	Proportion of chain retailers
propIndependent	Proportion of independent retailers
pctCloneTown	Proportion of 'clone' retailers
propVacant	Proportion of vacant retailers
propStructuralVacant	Proportion of vacant retailers since 2017
propVacantChange	Change in vacancy from 2017 - 2020
propComparison	Proportion of comparison retailers
propConvenience	Proportion of convenience retailers
propService	Proportion of service retailers
propLeisure	Proportion of leisure retailers
onlineExposure	Online exposure score (Singleton et al., 2016)
vulnerabilityIndex	Vulnerability index (Singleton et al., 2016)
eResilience	Composite e-resilience index (Singleton et al., 2016)
AvgIMDScore	Average IMD score of walking catchment
IMDDecile	Corresponding (national) decile for average IMD score

Finally, in Section 5.4 the role of the *structural* characteristics of retail centres in determining their response during this time is unpacked, through implementation of a modelling framework to quantify the impacts of different *structural* and catchment characteristics on changes to activity during this time, as well as considering how retail centre type (*function*) and region are related to such trends. In particular, the relationship between these independent variables and the change in share of total devices (i.e., activity) from a baseline (August-September average) to summer 2022 (June-July average) is modelled. Thus, for every retail centre the change in activity from 2021-2022 (Δ_i) is set as the dependent variable, and the *functional,* regional and *structural* attributes of the retail centre as the independent variables, as outlined in Equations 2 and 3 below.

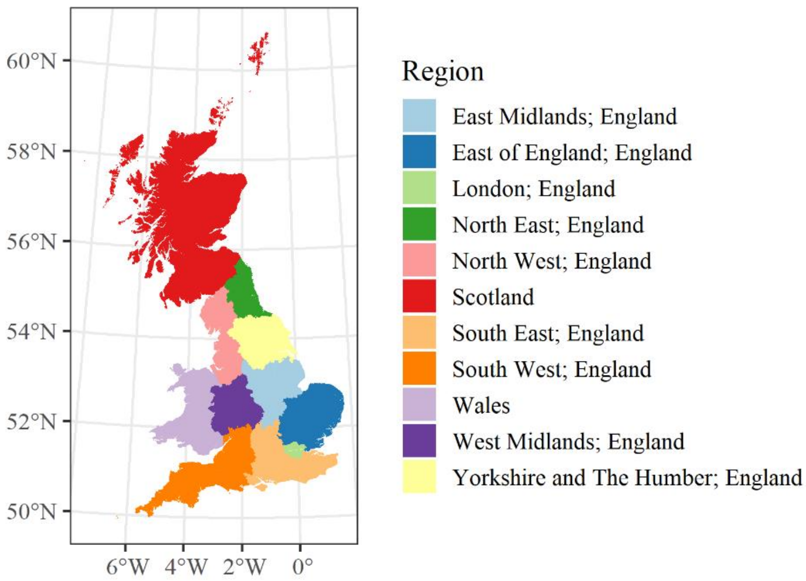

Figure 21. UK Regions, excluding Northern Ireland for the purposes of this study, as no retail centre indicators are available for retail centres in Northern Ireland.

$$\Delta_i = \beta_0 + \beta_1 \dots \beta_9 + \varepsilon$$

Equation 2. Model specification for *structural* (and catchment) characteristics of retail centres, following collinearity assessment of all variables in Table 10 (see Section 5.4).

$$\Delta_i = \beta_0 + \beta_1 \dots \beta_9 + \beta_{10} + \beta_{11} + \varepsilon$$

Equation 3**.** Model specification for *structural*, *functional* and regional characteristics of retail centres. Reference categories for β_{10} and β_{11} were Local Centres and Yorkshire and The Humber, due to low variation below in Section 5.4.

Where:

Δ_i = change in share of total devices between all retail centres nationally (%) from Aug/Sept 2021 to June/July 2022 for retail centre *i* (*continuous*).
β_1 = pctCloneTown (*continuous*).
β_2 = propVacant (*continuous*).
β_3 = propVacantChange (*continuous*).
β_4 = propComparison (*continuous*).
β_5 = propConvenience (*continuous*).
β_6 = propLeisure (*continuous*).
β_7 = propService (*continuous*).
β_8 = eResilience (*continuous*).
β_9 = AvgIMDScore (*continuous*).
β_{10} = *function* of retail centre *i* (*ordinal*).
β_{11} = *region* that retail centre *i* is located in (*nominal*).

5.4. Findings

The utility of Geolytix mobility data

As discussed in Section 5.3, significant attention must be paid to the representativeness and temporal stability of mobility data when seeking to explore temporal trends. Following direct consultation with Geolytix, it was suggested that their mobility data exhibits significant stability across days and weeks and between different regions, retail centre *functions* and directly comparable retail centres, despite a falling number of devices across the entire sample. To validate this and ensure this analysis did not fail to account for the changing number of devices, the proportion of national devices allocated to individual *functions* and regions at the weekly level to smooth variation in daily trends (Figures 22 and 23) was calculated and visualised, helping to identify whether robust comparisons could be made between retail centres, despite changing devices in the sample.

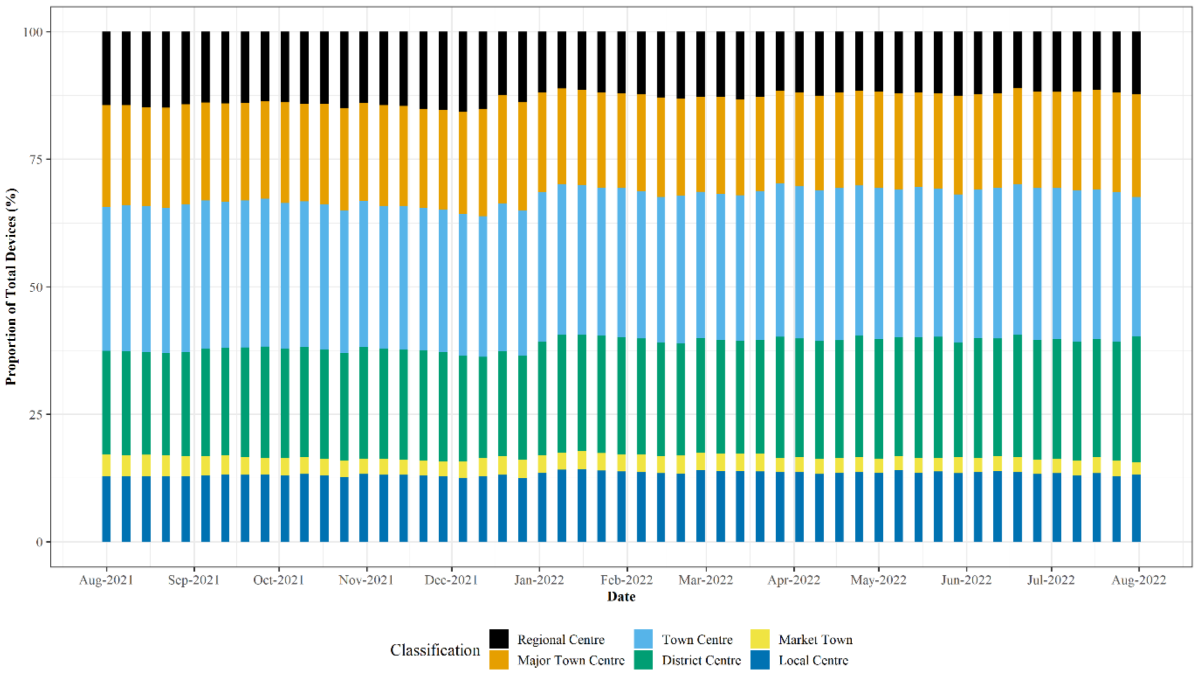

Figure 22. Stability of devices between different retail centre functions, highlighting the consistent share of devices between the six retail centre types over the study period.

From Figure 22 it is apparent that in terms of retail centre type (i.e., *function*), whilst the total number of devices in the sample fell dramatically over the study period (Figure 20A), the proportion of devices in each of the six types of retail centre remained largely consistent over the study period. This highlights that the mobility data does not bias certain types of centres, providing justification for comparison of share change in activity between different *functions* over time. In contrast, Figure 23 clearly illustrates that the loss of devices over the study period had a very distinct geography; it appeared to create a significant bias in London, where centres occupied a greater share of total devices nationally. However, it is not certain why this is occurring, as it could relate to movement of people back into London following the pandemic, or the growing popularity or accessibility to certain mobile phone providers in London which are unknown, so this must be controlled for. Thus, recovery trajectories for all retail centres outside of London are examined, as these trends are not subject to the inherent biases created by changing numbers of devices, resulting in a final sample of 862 retail centres. Whilst retail centres in London might comprise a more stable sample, the other ten regions experienced a consistent decline in the number of devices, so retail centres within these regions are directly comparable to each other, providing new insights, as opposed to existing literature on the response of retail centres to COVID-19 in London (Trasberg and Cheshire, 2021). However, what is not possible is exploration of individual retail centre trends over time, as they will be affected by the changing number of devices in the sample. Instead, comparison of retail centres within certain *functions* or in certain regions is more effective, as they have not been directly biased by this change in underlying devices, once London has been controlled for.

Exploring the response of retail centres

The response of different types of retail centres across the study period, as seen below in Figure 24, was of great interest. Firstly, there appeared to be no direct response to the arrival of Omicron in late November 2021 or its subvariants in February and May 2022, with the overall share of total devices between the six types of retail centre remaining largely unchanged in response to those key dates. This suggests that Omicron did very little to abruptly change the types of places people chose to shop, a direct contrast to what has been seen in earlier phases of the COVID-19 pandemic (Harris, 2022; Enoch et al., 2022; Ballantyne et al., 2022a; Frago, 2021). However, across the entire study period, there were interesting shifts in the change of share between the retail centre types, which raise significant questions about the longer-term recovery of different retail *functions*.

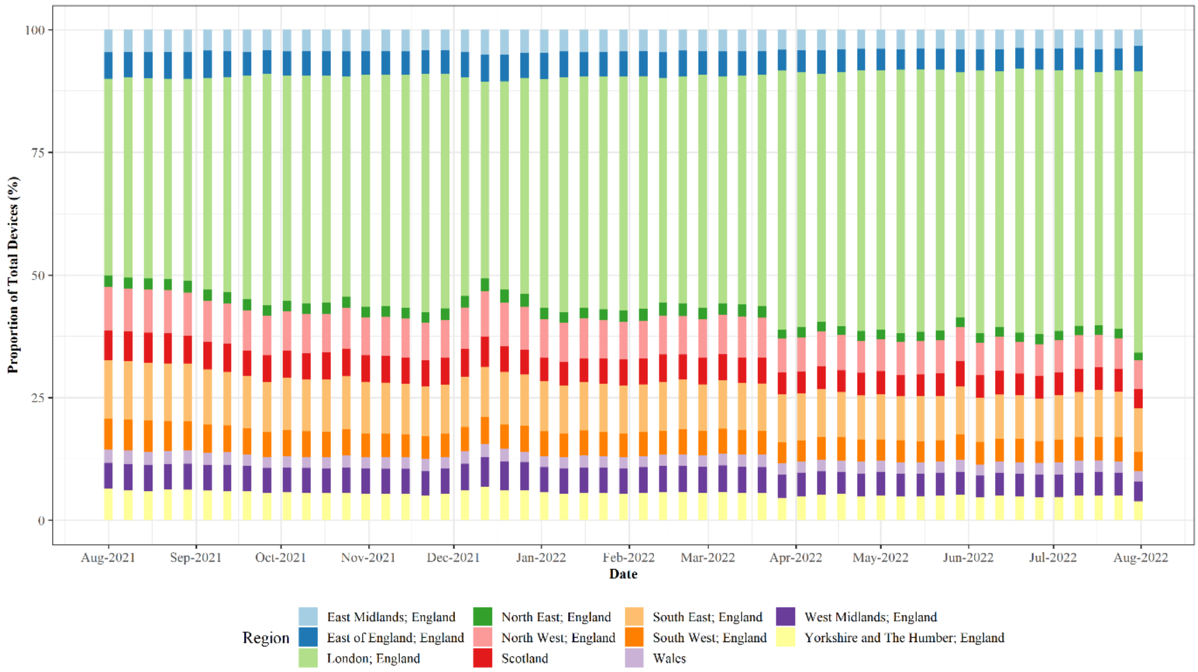

Figure 23. Stability of devices between different UK regions, highlighting the increasing share of devices in retail centres located in London over the study period, and general stability of device decline in all other regions.

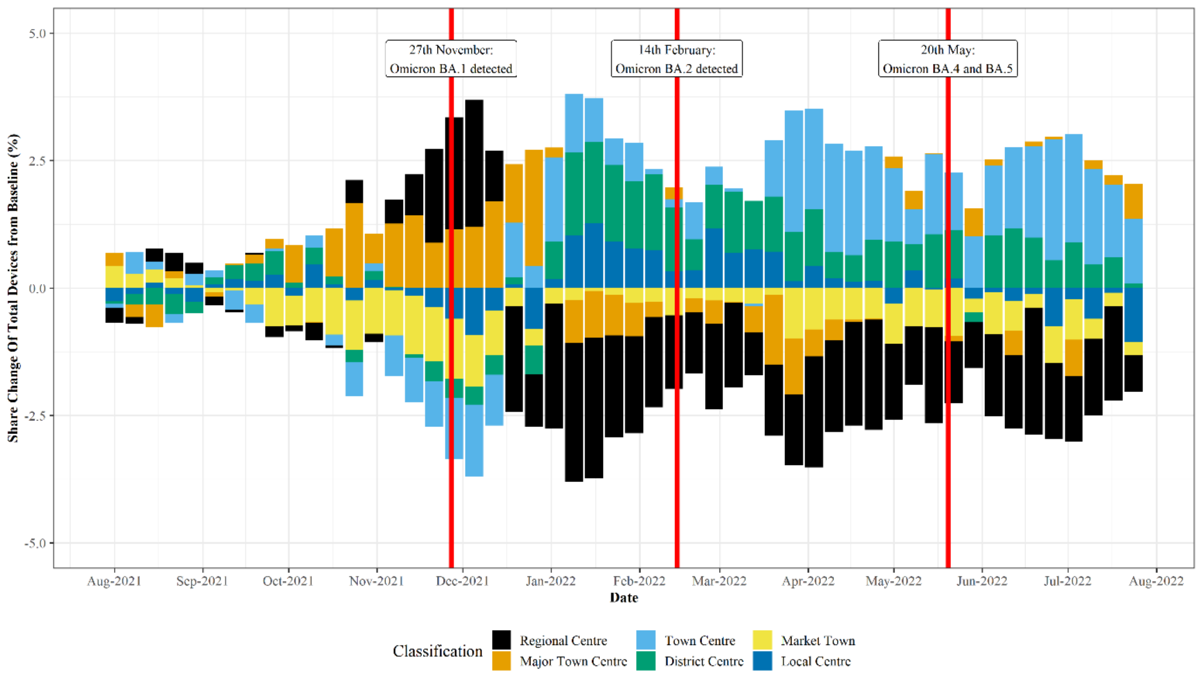

Figure 24. The functional response of retail centres visualised as the change in share of total devices (%) from the baseline, defined as the average share of devices (%) by retail centre type in August and September 2021.

For example, looking specifically at Regional Centres, the largest in size and typically the most diverse in retail offer (Macdonald et al., 2022), they exhibited a significant increase in share in the lead up to Christmas 2021, an expected trend given that these centres comprise the largest variety of retailers, products and ancillary activities, better fulfilling the needs of Christmas shoppers. However, what is most interesting is that following Christmas 2021, Regional Centres exhibited the most notable decline in share of activity from the baseline, suggesting that this specific *function* of retail centre has become less popular over the last year, relative to other retail centre *functions*, mirroring much of the literature seen earlier in the pandemic (Ballantyne et al., 2022a; Frago, 2021). Whilst this trend could be a result of shifts in consumer behaviour in response to Omicron or the recent cost-of-living crisis, it is certain that this trend is robust and not a product of falling devices in the Geolytix sample, given the examination of the stability of the dataset between different retail centre *functions* earlier in Section 5.4. On the other hand, Town Centres saw a reversal of share following Christmas 2021, where their importance became more significant following the Christmas period, similar to District Centres and Local Centres. These trends are interesting as during the first half of 2022, the UK was under "Plan B" restrictions, which were implemented to control the spread of the virus. Whilst not certain, it is not implausible to suggest that increasing activity in smaller retail centres (e.g., Local Centres) following Christmas and during 2022 was a result of risk-mitigation behaviours aiming to reduce exposure to Omicron during this time, as formal restrictions on mobility were not in place under "Plan B". This links to literature from earlier phases of the pandemic, where those *functions* deemed to be lower risk through a more 'localised' *function*, were those to experience the least significant impacts during the early stages of COVID-19 (Enoch et al., 2022; Frago, 2021; HSTF, 2021).

Similar trends can be seen when examining the recovery of retail centres in different regions too (Figure 25), which was posited to be a strong determinant of the economic performance of retail centres (Dolega and Lord, 2020). The largest decreases in activity were seen for retail centres in the South, specifically the South East and West, with noticeable decreases also seen in the North West and in Scotland. On the other hand, retail centres in East Anglia, East of England and West Midlands all appeared to experience significant uplifts in activity, when compared against the baseline period. Thus, what remains clear from this section is that *functionally* and regionally, there are significant disparities in terms of the recovery of retail centres during this time, with significant inequalities in how these recovery trajectories are manifesting between retail centres. Such inequalities are however not fully understood following exploratory analysis, as the responses of retail centres have been generalised based on *functional* and regional averages, instead of exploring individual responses.

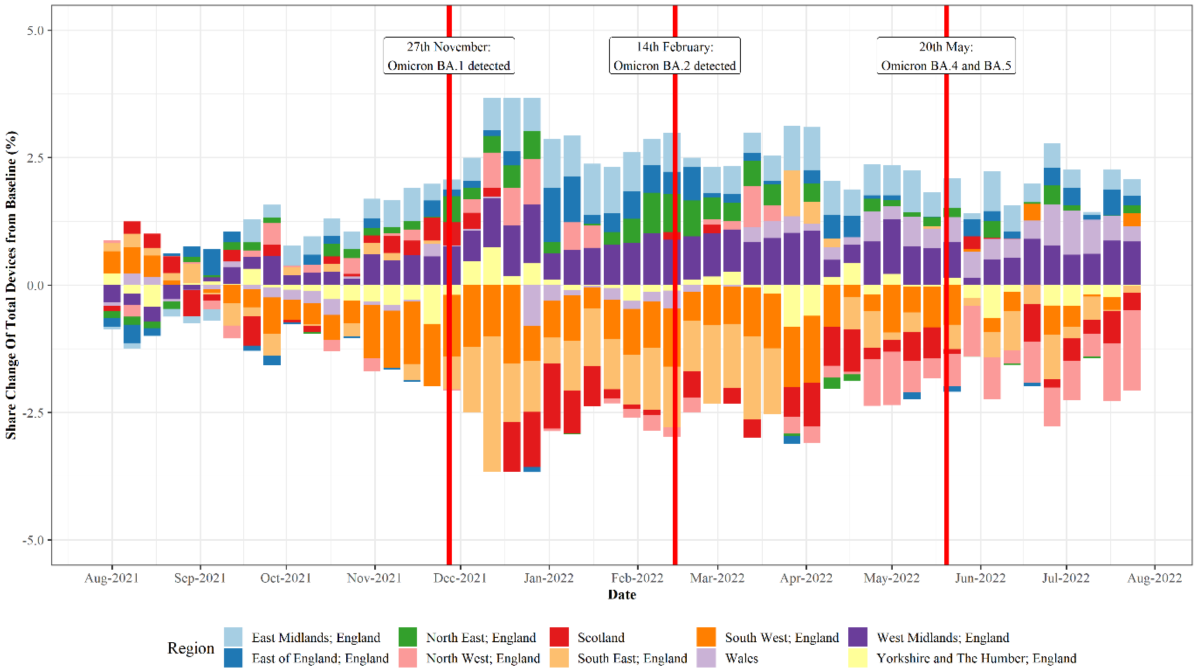

Figure 25. The regional response of retail centres visualised as the change in share of total devices (%) from the baseline, defined as the average share of devices (%) by region in August and September 2021.

Thus, to demonstrate the importance of considering these trajectories at greater resolution, below the individual responses of all Major Town Centres in the North-West (Figure 26A) and District Centres in the East of England (26B) are examined, highlighting the heterogeneity of responses between retail centres with the same *function* and regional geography. As above in Figure 24, Major Town Centres at the national level appeared to be experiencing an overall period of decline as opposed to recovery when compared against other types of retail centre in the UK, which theoretically should be more dramatic for those in the North-West of England (Figure 25). However, what is apparent from Figure 26 is that there is significant variation between retail centres, and whilst the majority did experience decline over the study period, though to varying degrees, there were some retail centres that experienced growth. Similarly, when looking at District Centres in East Anglia, whilst the majority are experiencing growth, though to varying degrees, there are still numerous retail centres experiencing decline, contrary to the national-level trends identified in Figures 24 and 25. Thus, this highlights the complexity of retail centre performance and recovery (Parker et al., 2017), which can be generalised to the national-level to provide a general overview of the role of *functional* and regional characteristics. However, significant variations in recovery clearly exist between individual retail centres that share similar characteristics, requiring analysis at a higher resolution to unpack some of these ideas. In addition, whilst *function* and region clearly interact with these trajectories, it is likely that the intrinsic *structural* composition of retail centres and their relationship to the catchment have a role too, as discussed in Section 5.2. Thus, an approach that can quantify these interactions more effectively is required, specifically one that can identify the relationships between *function*, region and *structure* on the trajectories of retail centre recovery, and quantify the importance of each.

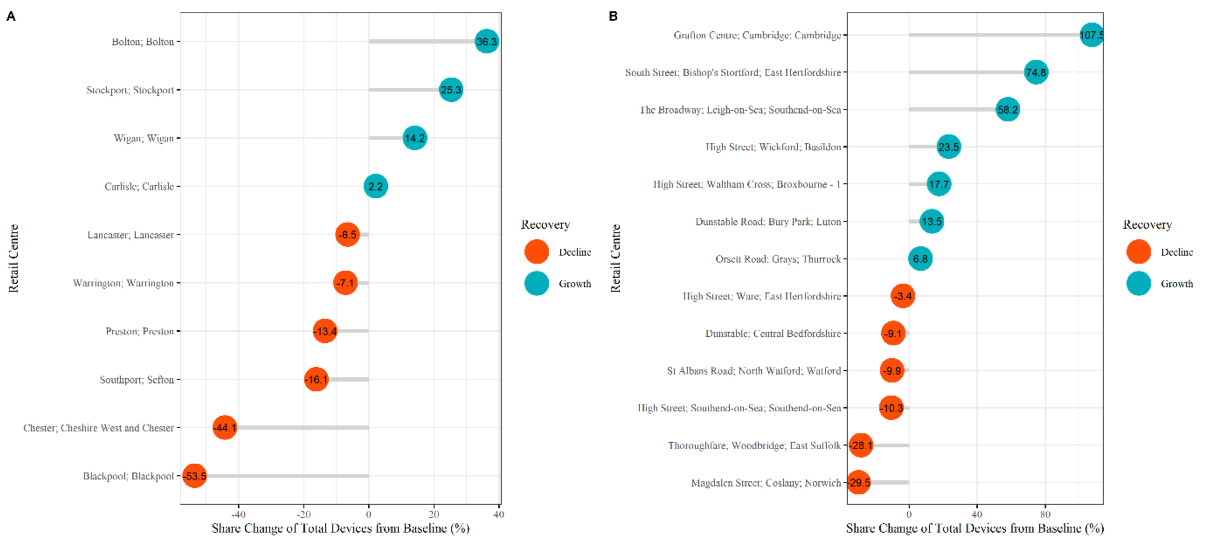

Figure 26. Recovery trajectories of Major Town Centres in the North-West (A) and District Centres in East of England (B). Trajectories have been calculated as the change in share of total devices (%) from the baseline, defined as the average share of devices (%) by region in August and September 2021.

Modelling the response of retail centres

As above, further analysis is required to unpack the significant amount of variation seen between the recovery of individual retail centres. Thus, in this section a modelling framework is deployed, as described above in Section 5.3, to quantify the role of *function*, region and the *structural* characteristics of retail centres in determining their response over the study period. Firstly, the prevalence of high collinearity between independent variables is examined utilising correlation analysis. Highly collinear variables were identified based on two criteria; those which have been used to create another independent variable such as onlineExposure and vulnerabilityIndex which are put together to construct eResilience, or those where the correlation coefficient exceeded 0.7 or was lower than -0.7. Following removal of *structural* characteristics with high collinearity, a model was fit (see Equation 2) to first assess the role of the *structural* (and catchment) characteristics of retail centres in determining the change in activity between summer 2021 and summer 2022 (Δ_i), as described above in Section 5.3.

The results of a model fit with just the *structural* characteristics can be seen below in Table 11, where coefficients are interpreted as the estimated percentage change in retail centre activity (share of total devices) given a one-unit change in each of the explanatory variables. The results suggest that in general, the *structural* (and catchment) characteristics of retail centres are associated with Δ_i, though to varying degrees. For instance, those with higher proportions of Leisure retailers were more likely to experience negative growth (-0.735), as at the beginning of the COVID-19 pandemic (Enoch et al., 2022), whilst those with higher proportions of Service retailers were more likely to experience growth (0.830). This suggests that over the 12-month study period, retailers with a more 'essential' retail offering were those that occupied a greater share of consumers, which is supported by a positive coefficient for propConvenience and negative coefficient for propComparison, though both were not statistically significant. Whilst both statistically insignificant, the coefficients for variables describing the vacancy of retail centres were also of great interest; both exhibited negative coefficients suggesting that retail centres struggling with larger numbers of empty stores typically experienced negative growth during the study period, a well-documented determinant and consequence of the changing economic performance of retail centres (Dolega and Lord, 2020; Enoch et al., 2022).

Table 11. Model results for *structural* (and catchment) characteristics of retail centres.

Variable	**Coefficient**	**p value**	**Sig.**
pctCloneTown	0.151	0.390	-
propVacant	-0.304	0.523	-
propVacantChange	-0.158	0.764	-
propComparison	-0.433	0.167	-
propConvenience	0.467	0.448	-
propLeisure	-0.735	0.039	*
propService	0.830	0.018	*
eResilience	0.369	0.030	*
AvgIMDScore	0.429	0.011	*

Significance levels: < 0.05 *, < 0.01 **, < 0.001 ***, R-squared: 0.33, Adjusted R-squared: 0.23.

Unsurprisingly, the resilience of retail centres to online shopping (eResilience) was seen to have a positive effect on the recovery of retail centres during this time; retail centres with a high resilience to online shopping came to occupy a greater share of consumers between 2021 and 2022. This is interesting, as in the UK the e-resilience of centres has long been considered a vital determinant of their economic performance both in and out of the COVID-19 pandemic (Singleton et al., 2016; Enoch et al., 2022), and appears to still be a key factor. This raises interesting debates about the continued plurality of different retail centres; those deemed to provide an offering that will not be overshadowed by online shopping (i.e., higher e-resilience), have recovered faster and appear to be maintaining such recovery, when compared with those more susceptible to the effects of 'E-commerce'. This is a similar trend to what was seen in earlier phases of the pandemic, where large numbers of people were switching to online purchasing (Ntounis et al. 2020), only visiting stores/retail centres where they could access a good or service less suited to 'E-commerce', typically service and/or convenience retailers, both of which exhibited positive coefficients above in Table 11. From a conceptual standpoint, this is interesting as this measure accounts for the structural components and level of 'supply' (Singleton et al., 2016), but also incorporates catchment characteristics through quantification of the 'online exposure' of the catchment (i.e., demand), demonstrating the importance of understanding the role of supply and demand when trying to unpack the response of retail centres, and their economic performance, as in Jones et al. (2022).

The final independent variable that exhibited a statistically significant association with change in activity was deprivation (AvgIMDScore), as initially suggested by Dolega and Lord (2020), where retail centres in more deprived areas were seen to occupy a greater share of consumers, i.e., recovering at a faster rate. This is an interesting finding, and the first to link the economic performance of retail centres directly to the deprivation of its catchment. A plausible explanation could relate to the implementation of "Plan B" recommendations, which occurred during the study period (November 2021 – February 2022) to reduce the spread of Omicron. It is well documented that neighbourhoods with differing socio-economic and demographic showed different levels of engagement with government restrictions and vaccination programmes throughout the pandemic (HM Government, 2022). This could be apparent here, where people in more deprived areas could have been less likely to follow to government recommendations and reduce their mobility during this time, resulting in higher activity in nearby retail centres, as above in Table 11.

Thus, some interesting associations between the *structural* characteristics of retail centres and their recovery trajectories over the study period have been identified. However, given exploration of the response of retail centres with different *functions* and regional geography in Section 5.4, it is important to incorporate such insights into the modelling framework, to identify the concurrent role of *function,* region and *structure* in determining the response of retail centres. The results of the model with only the significant *structural* indicators from Table 11, and dummy variables for retail centre *function* and region can be seen below in Table 12 (see Equation 3). The coefficients for region and *function* can be interpreted as the average change in retail centre activity for the comparison group relative to the reference group, keeping all other variables constant. The reference groups were selected as Local Centres (Classification) and Yorkshire and The Humber (Region), given their low variance across the study period, as identified in the previous sub-section.

Table 12. Model results for the *structural*, *functional* and regional characteristics of the retail centres. Reference categories for Classification and Region are 'Local Centres' and 'Yorkshire and The Humber' respectively.

Variable	**Coefficient**	**p value**	**Sig.**
propLeisure	-0.537	0.089	-
propService	0.968	0.001	*
eResilience	0.348	0.028	*
AvgIMDScore	0.164	0.050	*
(Classification) Regional Centre: Local Centre	-2.810	0.832	-
(Classification) Major Town Centre: Local Centre	-1.920	0.757	-
(Classification) Town Centre: Local Centre	2.620	0.510	-
(Classification) District Centre: Local Centre	4.580	0.320	-
(Classification) Market Town: Local Centre	-9.520	0.049	*
(Region) East Midlands: Yorkshire and The Humber	1.400	0.843	-
(Region) East of England: Yorkshire and The Humber	-6.960	0.303	-
(Region) North East: Yorkshire and The Humber	-1.270	0.886	-
(Region) North West: Yorkshire and The Humber	0.959	0.877	-
(Region) Scotland: Yorkshire and The Humber	-8.580	0.275	-
(Region) Wales: Yorkshire and The Humber	3.333	0.651	-
(Region) South West: Yorkshire and The Humber	-15.200	0.018	*
(Region) South East: Yorkshire and The Humber	-10.800	0.078	-
(Region) West Midlands: Yorkshire and The Humber	3.310	0.621	-

Significance levels: < 0.05 *, < 0.01 **, < 0.001 ***, R-squared 0.30, Adjusted R-squared: 0.25.

Similar to the earlier discussion, propService, eResilience and AvgIMDScore exhibited statistically significant positive associations with retail centre activity, which can be interpreted as increasing the overall recovery of retail centres during this time. In terms of the *function* of retail centres, the direction of the coefficients aligned with earlier findings about the recovery or growth of retail centres during this period; for example, retail centres classified as Major Town Centres, Regional Centres and Market Towns were all found to have negative associations with Δ_i on average, relative to Local Centres, as in Figure 24. In comparison, District Centres and Town Centres exhibited positive associations, again matching the discussion earlier. However, it is important to consider these findings in relation to their statistical significance; very few retail centre *functions* exhibited statistically significant associations with the change in retail centre activity between 2021 and 2022; Market Towns were the only retail centres to exhibit a statistically significant relationship with Δ_i. Whilst the use of share change (over total devices) could flatten the significance of *functional* differences in recovery, it appears that *functional* differences are of less significance than the *structural* and catchment characteristics of retail centres in determining recovery, an interesting finding.

Similarly, when looking at responses between regions (Table 12), the direction of the coefficients was again unsurprising, with those regions identified in decline earlier (Figure 25) such as the South East, South West and Scotland all having negative coefficients, relative to Yorkshire and The Humber, though not all were statistically significant. What is particularly interesting is that the region that appeared to experience some of the most significant reductions in share in Figure 25, the South West, had a statistically significant negative association with retail centre activity, detailing that retail centres in the South West were more likely to experience decline than recovery during this period, relative to the reference category and keeping all other indicators constant. However, as with *functional* responses, it is important to reiterate that most regions exhibited statistically insignificant relationships with Δ_i during the study period.

Thus, what remains clear from this modelling exercise is that retail centre recovery (Δ_i) during this time is dependent on the overall *structure*, *function* and regional geography of the retail centres, though to varying degrees, with *function* and regional geography contributing significantly less. It appears that the *structural* and catchment characteristics of retail centres remain a greater determinant of changes to retail centre activity during this time, thus more research is needed to unpack how at finer geographical resolutions (as opposed to regions), different *structural* characteristics of retail centres geographies determine such responses

(Dolega and Lord, 2020; Philp et al., 2022). However, there are lots of additional unanswered questions that need addressing, such as the role of multidimensional typologies (e.g., Dolega et al., 2021), seasonal and weather effects (e.g., Rose and Dolega, 2022) and the recent cost-of-living crisis, which has exacerbated inequalities between different regions (Wood, 2019). Furthermore, it would be of great utility to identify how and when these recovery trajectories began, given data with a longer timescale, though this was not possible with the Geolytix data used in this investigation.

5.5. Discussion and Conclusions

Spaces of consumption such as retail centres have faced significant challenges in recent years, with the COVID-19 pandemic continuing to exacerbate the decline of physical retail spaces. Whilst some studies have explored the response of consumption spaces to the pandemic, they are often restricted to specific geographic areas, or tend to focus on the impacts of national lockdowns during the earlier waves of the pandemic. Using mobility data from Geolytix, the recovery of retail centres across Great Britain was investigated, during a period characterised by the Omicron variant. These findings are of great significance, providing an overview of the response of retail centres at the national level for the first time, demonstrating that such responses were partially determined by the *functional, structural* and regional characteristics of the centres.

Perhaps the most important finding was that the response (and recovery) of retail centres was not homogenous, providing evidence that examination of national trends of retail centre recovery, as in Section 5.4, are not enough to capture variation in responses between a network of centres with different *functional,* regional and *structural* characteristics. By modelling the nature of these recovery trajectories between centres with different characteristics in Section 4.3, it is clear that there were specific ‘winners’ and ‘losers’ during the study period. Functionally, whilst retail centres towards the top of the hierarchy (e.g., Regional Centres) appeared to exhibit the most pronounced recovery leading up to Christmas 2021, this trends reversed in 2022, where the popularity of retail centres at the cores of major towns and cities saw decline rather than growth, as earlier in the pandemic (Ballantyne et al., 2022a; Frago, 2021). In addition, significant regional inequalities in retail centre recovery were identified, such as the apparent decline of retail centres in the South (excluding London), whilst retail centres in the Midlands, Wales and areas of the North exhibited the opposite trend.

Finally, specific *structural* characteristics that were associated with stronger recovery were identified; lower dominance of 'non-essential' retail (e.g., Leisure), higher resilience to online shopping and greater levels of deprivation within the catchment, with *structural* characteristics appearing to be a greater determinant of recovery than the overall *function* or regional geography of retail centres.

However it is important to remain cautious of these trends, especially given they are based on exploratory analysis and modelling, which did not account directly for the impacts of seasonality, weather and holiday periods (Lyu et al., 2022; Rose and Dolega, 2022), and based on trends for a subset of the major retail centres across the UK. Further research should seek to identify what additional knowledge can be generated about retail centre recovery by focusing on retail centres in London, or those 'Small Local Centres', which comprise the largest proportion of retail centres in the UK (Macdonald et al., 2022). However, perhaps the greatest consideration relates to the underlying limitations of the mobility data used in this study. Mobility data often has a tendency to introduce generational and/or spatial biases, as previously identified (Trasberg and Cheshire, 2021), but it is also important to think critically about the temporal stability of the dataset, which as a result of significant reductions in the number of devices and applications over time (Figure 20A), curtailed exploration of individual recovery trajectories over time, instead resulting in comparisons between similar areas and modelling of change in the share of activity between two time periods. As a result, there remains significant uncertainty as to the exact nature of retail centre recovery, a major challenge when trying to utilise 'Big Data' in Urban Informatics (Shi et al., 2022).

However, given significant effort devoted to controlling for the temporal instability of the dataset, through identification of relative stability between all retail centre *functions* and most regions (see Section 5.4), the findings presented are empirically robust. Whilst there are some important considerations to make about the temporal stability of such data before using it to answer new research questions, correct use of mobility data offers significant advantages over other economic performance measures for retail centres. For example, mobility data does not privilege certain geographic areas or locations within retail centres, as is the case with footfall sensors (Philp et al., 2022), and typically offers a greater temporal resolution than other 'static' measures of economic performance, such as vacancy rates (Dolega and Lord, 2020). However, it would still be more preferable to use actual sales data to monetise the performance of retail centres, as is the case with individual stores (e.g., Rose and Dolega, 2022), but the potential to do so has not yet been realised, given a lack of suitable data.

To conclude then, the results of this study provide empirical evidence of the recent recovery of retail centres, highlighting that there are certain *functional*, regional and *structural* characteristics associated with particularly stronger recovery trajectories, contributing further to the narrative that retail centre performance is multidimensional (Parker et al., 2017). In this sense, it is argued that national policies seeking to maintain or improve the vitality or viability of consumption spaces need to account for this added knowledge. By considering the *functional* role of the retail centre and its *structural* and catchment characteristics, and constructing a 'Digital Twin' framework, researchers can use advancements in Big Data and modelling to simulate how such policies can result in positive outcomes for consumption spaces (Goodchild, 2022; Shi et al., 2022). Such interventions have never been more important, as whilst the COVID-19 pandemic remains present, the retail sector is also subject to the recent cost-of-living crisis, where its impacts are already apparent in falling sales and footfall in recent months (ONS, 2022; Wright, 2022). Given the rising costs of energy and food, increasing taxes and wages falling in line with increasing inflation in the UK (Patrick and Pybus, 2022), the retail sector is expected to continue to face some of the most significant impacts, with falls in consumer confidence and a new wave of retail vacancies expected in the near future. This raises significant questions, which are not new, but remain important about the trajectories of retail centre performance in the near future, and the social and economic value that these urban phenomena represent. These issues are however not well understood, and there is a broader agenda for further research into the continued monitoring of retail recovery and decline, utilising retail centre geographies as geographic data tools to provide evidence that can inform policy decisions and provide solutions to both acute and longer-term issues. This study provides an initial basis upon which to do so, through examination of national-level trends in retail centre activity, utilising unstable data derived from mobile phone applications.

Acknowledgements

The authors would like to thank Geolytix and Maurizio Gibin (CDRC) for facilitating access to the data, and to Christoph Mulligan (Geolytix) for their advice in better handling the temporal instability of the data.

6. Thesis Discussion and Conclusions

This chapter concludes this thesis by providing a summary of its key findings (6.1), before discussing its key contributions and implications (6.2), limitations (6.3) and arising areas of future research both in retail geographies and other fields (6.4).

6.1. Research Findings

Across the three empirical chapters, this thesis has addressed its research aims (Section 1.3). The following narrative re-introduces these aims in the context of the relevant chapters and findings, to outline the ways in which the aims have been fulfilled.

Aim One: Investigate whether recent advances in retail centre delineation and classification can be used to capture the geographies of retail centres into other international settings.

Chapter three fulfilled the first aim by investigating the potential for existing approaches to yield understandings about the geographies of retail centres in other international settings. Notable findings are presented about retail centre geographies in Chicago and the appropriateness of such methodologies for obtaining such insights. Using data from SafeGraph, 1,599 retail centres were delineated across the MSA, and their functional ecologies were represented as a 'two-tier' non-hierarchical retail centre typology. Through statistical validation and exploration of their response to the COVID-19 pandemic, it is argued that the retail centre geographies presented here are robust, and of great utility, constructed using cutting-edge techniques in retail centre delineation and classification. However, whilst chapter three represents the first set of retail centre geographies for the U.S., the emphasis was in unpacking the potential of recent advancements in retail centre delineation and classification (aim one), something that has arguably been achieved in the third chapter. Firstly, the modified-DBSCAN approach (Pavlis et al., 2018) was deemed unsuitable for the study area and dataset, thus an approach using HDBSCAN, network distance matrices and H3 was presented, resulting in a delineation that is both robust, and simpler to implement and understand, although spatially constrained to Metropolitan areas, with further work needed to identify methods suitable for larger spatial scales. Secondly, the classification framework (Dolega et al., 2021) was of great utility, enabling construction of a retail centre typology that better represents the spatiality of retailing. However, further enhancement and customisation is required to capture specific niches in U.S. retail, and to ensure understandings of the geographies of retail centres are fit-for-purpose. Thus, to return to aim one, a 'data-driven' analytical framework, through which to develop such understandings, has been established, helping to generate new knowledge about retail centre geographies in new locations; *where* they are located and *what* characteristics they have.

Aim Two: Generate a comprehensive understanding of the geographies of a national retail centre system outside of the UK.

Chapter four fulfilled the second aim, by utilising the insights obtained in chapter three to generate a comprehensive understanding of retail centre geographies for the U.S. Specifically, in terms of *where* retail centres are located, using a new retail centre delineation method based around H3. 10,956 major retail centres were identified across the U.S., typically in the most urbanised areas, varying in scale and form, and were deemed to be much more accurate than those in chapter three, through incorporation of additional geographic information about retail location (e.g., building footprints, land-use polygons). The functional ecologies of the retail centres were again represented as a 'two-tier' typology (the *what*), but at a much finer resolution, owing to a greater breadth and depth of input variables, and customisation of the original framework to capture specific niches in American retailing. Finally, a series of catchments were extracted which provide information about *who* is using these retail centres and where they come from, through calibration of a Huff model with mobility data, and using it to estimate retail centre catchments. Thus, chapter four arguably fulfils aim two by providing a comprehensive overview of the *who, what* and *where* of U.S. retail centre geographies for the first time. Whilst chapter four presents these as individual geographies, it also contributes by formalising the apparent connections between them in a new conceptual framework. This framework provides for the first time, a conceptual understanding of the interactions between retail centre geographies in a network of retail centres at the national level. In particular, it highlights the connections between the *where, what* and *who*, such as the evident interactions between scale (the *where*) and function (the *what*), whilst also demonstrating that these geographies are better understood when considered together, for example, by supplementing information from the typology into catchment calibration. Thus, to return to aim two, a comprehensive understanding of the geographies of retail centres in the U.S. retail centre has been provided, which has shown that the system is heavily inter-connected, better studied through the use of the integrated empirical and conceptual frameworks that have been presented in chapter four.

Aim Three: Explore spatio-temporal trends of retail centre recovery using data derived from mobile phone applications.

Chapter five fulfils this aim by utilising a large, unstable mobility dataset to examine the recovery of retail centres between 2021 and 2022, demonstrating *how* retail centre geographies can be used as geographic data tools to understand the response of the retail sector. Specifically, using data from Geolytix, the chapter considers the recovery trajectories of British retail centres, exploring the change in share of activity between different *functions* and in different regions, before modelling individual recovery trajectories in relation to *function,* region and the *structural* characteristics of the retail centres. By exploring and modelling these recovery trajectories, significant heterogeneity was documented, reiterating the importance of multidimensionality when assessing the changing economic performance of retail centres. Specifically, *function* and region appear to be less associated with recovery, whilst it appears that the *structural* characteristics of the retail centres, particularly composition of leisure and service retailers, e-resilience and deprivation appear to be closely associated with processes of recovery and decline during this period. Furthermore, the chapter contributes significant insights about the utility of an unstable mobility dataset for examining such trends, highlighting in particular the value of such data when treated as snapshots rather than a time-series, and reiterating the importance of considering the underlying sample of devices used to build the dataset (and changes to it). However, it is important to reiterate that the primary aim of this chapter was to unpack the nature of these spatio-temporal responses, which has arguably been fulfilled. Thus, to return to aim three, a national overview of the recovery of retail centres in the UK using data from mobile phone applications has been provided, highlighting key considerations that are of greater significance when dealing with unstable mobility data, and that recovery trajectories of retail centres are not homogenous, dependent partially on the *function,* region, and the *structural* characteristics of retail centres.

6.2. Contributions and Implications

In seeking to address the three primary research aims, this thesis makes a number of significant contributions, specifically in the use of new data, application of new methodologies, establishment of new empirical and conceptual frameworks, contributing to key theoretical debates, and generating new knowledge about the geographies of retail centres both in the U.S. and in the UK. These contributions create significant implications which will be of great interest to lots of practitioners, including relevant stakeholders, town centre managers, retailers, and wider academia, providing an evidence base upon which to develop retail policy and future research proposals to unpack further some of these findings.

New sources and forms of data

This thesis has utilised a number of new data sources to generate understandings of retail centre geographies. Data from SafeGraph has been a key source of information for retailer locations in this thesis, representing its first use in deriving insights about (retail) agglomerations, despite the general consensus that it represents the best source of openly accessible information (for academics) about retailer locations in the U.S. In particular, the SafeGraph places database enabled identification of *where* retail centres are located and *what* characteristics they have, as in chapters three and four, and it is important to note that such insights would not have been as robust with other openly accessible data sources (e.g., OpenStreetMap), as SafeGraph offers better geographic coverage, is updated frequently and contains a large number of attributes and information on POIs. SafeGraph geometries were also used in chapter four to better capture the spatial extent of large retail developments, and their in-house mobility dataset 'patterns' was also used to derive proxy performance measures in chapters three and four, helping also to investigate the response of Chicago retail centres to the early weeks of the COVID-19 pandemic (Section 3.6). Thus, this thesis has highlighted the utility of SafeGraph data in the field of retail geography, which through the 'SafeGraph data for Academics' programme, provides researchers with access to a vast array of data for use in research, without which this PhD thesis and its contributions would not have been as significant. It is important to state that SafeGraph data might be inherently subject to limitations, especially given the 'black-box' nature of how the 'patterns' mobility data is collected, as well as a significant lack of detail about how individual POIs and building geometries are obtained. However, these were unavoidable limitations as SafeGraph provided such a vast quantity of high-resolution data on retailer locations that could not be obtained elsewhere freely, and as such the understandings of the *who, what* and *where* of U.S. retail centre geographies would not have been as robust without it.

Chapter five saw the use of new datasets to fulfil aim three. Firstly, a new suite of retail centre indicators, which were developed as part of this thesis project (CDRC, 2021b), were used to capture key *structural* differences between retail centres in the UK, providing a mechanism through which to explore how retail centres with different *structural* and catchment characteristics (e.g., deprivation) have recovered in recent months, representing the first use of this data in academic research. The use of this dataset contributes significant value, by demonstrating that formal indicators provide a mechanism through which to explore the changing nature of retail centres and their response to external phenomena, resulting in new insights about the determinants of retail centre recovery in chapter five. For example, it was identified that the resilience of retail centres to online shopping (e-resilience), deprivation of the catchment and structural composition of retail centres were closely associated with recovery and decline, a contribution that would not have been possible without the retail centre indicators (CDRC, 2021b).

Furthermore, in chapter five the Geolytix in-app aggregated location dataset (CDRC, 2021a) was used to provide a mechanism through which to measure the economic performance of retail centres, and examine their trajectories of recovery (or decline) over time. This dataset offered significant potential as proxy performance measure for retail centres, given the limitations of other proxy measures such as vacancy rates which are often static (Dolega and Lord, 2020) and footfall data, which typically suffers from poor geographic coverage and representation (Philp et al., 2022). Whilst the Geolytix data was subject to significant limitations as reviewed below in Section 6.3, it offered national coverage, is available at a high spatio-temporal resolution and can be easily appended to retail centre boundaries using H3, thus helping to fulfil the third aim of this thesis in chapter five, as well as creating significant potential for future research opportunities (Section 6.4), once its limitations have been controlled for effectively; a contribution made in this thesis.

Methodological approaches

This thesis has contributed significantly by providing a number of new methodological approaches and tools to the field of retail centre geographies. Firstly, in terms of retail centre boundary delineation, two new approaches were presented, the first utilising HDBSCAN, network distance matrices and H3 (chapter three), which was deemed to be an effective alternative to existing examples (Pavlis et al., 2018), better accounting for local point density heterogeneities, and accounting for local urban morphology through incorporation of precomputed network distance matrices. Chapter three also generated significant scope for the use of H3 as an important tool, as an effective and interpretable way of refining cluster boundaries, and as a result was used to delineate retail centre boundaries in chapter four, adapted from that in Macdonald et al. (2022). This approach, which resulted in retail centre boundaries for the national extent of the U.S., incorporated a wider range of spatial information about retail, including building footprints and land-use polygons, resulting in better capturing of large, purpose-built retail developments. Furthermore, the method was highly scalable and reproducible (see Ballantyne, 2022), enabling delineation of retail centre boundaries for the national extent of the U.S., a limitation of chapter three, and other existing approaches (Pavlis et al., 2018).

Throughout chapters three and four, significant enhancements and modifications to the original retail centre typology framework (Dolega et al., 2021), as well as extension into new international contexts with new data sources was demonstrated. In both chapters, the framework is applied to the U.S. for the first time, using data from SafeGraph and other ancillary datasets, a novel contribution in itself. However, in chapter three, minor modifications were made to suit available data, and in chapter four, building on the idea that specific 'niches' in American retail existed, the framework was significantly enhanced and customised, through capturing of a greater breadth of input variables and increasing the representativeness of classification domains, resulting in a more 'fit-for-purpose' U.S. retail centre typology, which provides an empirical basis upon which to extent its application further into other international settings or subjects of study (e.g., Shannon, 2016). By accruing new sources of data, considering the ways in which the original framework needs to be modified to suit the setting or subject of study, and utilising the code and GitHub repository associated with Chapter four (Ballantyne, 2022), multidimensional understandings and insights of retail (or other) environments can be gained that will be of great interest to lots of different fields.

A significant research gap identified in Section 1.2 was that approaches to generating retail centre catchments are sparse in the literature, especially using 'non-conventional' forms of data. However, in chapter four a new methodology was presented through which to estimate retail centre catchments for U.S. retail centres, using mobility data from SafeGraph. The Huff model for retail centre catchments (Dolega et al., 2016) was adopted, before modifying measurement of 'attractiveness' based on available data and new knowledge about the conceptualisation of 'attractiveness', and calibrating the model with SafeGraph patterns data.

The patterns data provides 'observed' patronage behaviours, and using advancements in model calibration techniques (e.g., Wang et al., 2016), models estimated 'predicted' patronage behaviours which accurately captured the 'observed' behaviours seen in the patterns data. This is a highly significant contribution, representing an entirely new approach to estimating catchments for retail centres, through the use of mobility data, which is openly available to academics, as opposed to sensitive commercial data, which has characteristically dominated such efforts in the past (e.g., Dolega et al., 2016). As a result, the new approach presented in chapter four represents a significant methodological contribution, and one that has generated a suite of empirically and statistically robust catchments through which to better understand *who* is using these retail centres and where they are coming from, using new forms of data.

Finally, in addressing aim three, new tools were presented for how to handle unstable data generated by mobile phone applications (chapter five). Through examination of the stability and representativeness of the Geolytix aggregated in-app location dataset in different study areas and *functions*, it became possible to identify a stable set of retail centres to study in relation to their recovery. Furthermore, in chapter five it was demonstrated that it is more useful to consider unstable mobility data as snapshots of areas rather than a time-series, following representativeness and stability analysis, enabling interesting comparisons to be drawn between areas. Thus, this represents a formal methodological contribution for two reasons. Firstly, chapter five demonstrates a new approach to handling unstable mobility data, where the sample of devices does not remain consistent over the study period, which can be of great utility given the use of such data in related studies (e.g., Calafiore et al., 2022), Furthermore, the use of such data in this way represents the first conceptualisation of retail centre performance using mobility data, providing an empirical basis upon which to continue to do so, once the limitations of such data are controlled for.

Theoretical and conceptual understandings of retail (centre) geographies

The thesis has contributed significantly to key theoretical and conceptual debates underpinning retail (centre) geographies. Firstly, further evidence of the need for post-hierarchical understandings of retail environments has been presented (chapters three and four), which can benefit both the academic community in efforts to conceptualise the organisation of urban agglomerations, but also practitioners and stakeholders who rely on hierarchical understandings, and could benefit from alternative approaches, given the dynamic and complex nature of retailing in the 21st century. By using variables deemed fundamental to understanding the contemporary landscape, classification based on similarity and salient characteristics, and customisation of the framework to capture specific niches (chapter four), a more representative insight into the spatiality of retailing has been provided (Dolega et al., 2021). In particular, it is now evident that non-hierarchical structures (and functions) exist in the system which do not fit conventional retail centre hierarchies underpinned by CPT. Thus, whilst hierarchical structures still have significant utility in terms of their ease of interpretation and familiarity to stakeholders, practitioners and policy makers, as well as low cost in construction owing to their general simplicity, it is argued that a national system of retail centres cannot be summarised based solely on differences in supply/demand, as the contemporary retail environment is much more complicated, calling for greater multidimensionality in its conceptualisation.

The second theoretical contribution relates to the Huff model, which has historically occupied a place of great significance in retail geography. Chapter four provides significant evidence that the Huff model has retained its conceptual relevance, owing to advancements in the way these models can be calibrated with new forms of data. Whilst this clearly demonstrates the continued utility of the Huff model by generating robust insights about *who* uses retail centres, it also yielded substantial theoretical insights about the drivers of demand in the contemporary retail environment; retail centre attractiveness remains less important than distance, with this effect becoming much stronger for those functions specialising in an 'everyday' retail offering. Furthermore, chapter four demonstrates that measures of attractiveness need to be more multidimensional, considering a wider pool of factors beyond just the size of the retail or shopping centre (Dolega et al., 2016; Gong et al., 2021). Therefore, the theoretical underpinnings upon which the Huff model is based remain of significance in the contemporary retail environment, given the availability of new forms of data and conceptual advancements, despite major shifts in the way retailing is organised.

The final theoretical contribution of this thesis relates to the way in which other existing conceptual and empirical frameworks have been built upon, bringing them together to better understand retail centre geographies within a multi-national context. Thus far, literature on retail centre geographies has typically focused on singular geographies, considering in isolation *where* they are located (e.g., Thurstain-Goodwin and Unwin, 2000; Pavlis et al., 2018; Macdonald et al., 2022), *what* characteristics they have (e.g., Brown, 1992; Dolega et al., 2021) or *who* uses them (e.g., Dolega et al., 2016; Lloyd and Cheshire, 2017; Jones et al., 2022). Whilst different methods were applied for each, chapter four demonstrates that they are connected through the use of individual geographies to extract others; for example, the use of the typology in catchment calibration to improving the robustness of estimates. Furthermore, significant connections were evidenced between these geographies, such as the mapping of retail functions across spatial scales, and the importance of function in determining patronage. However, thus far there has been no consensus on how best to bring these three geographies together to provide a comprehensive overview of a national retail (centre) system. Thus, this thesis has established a new conceptual framework, which argues that understandings of retail centre geographies are more comprehensive and useful when considering the *who, what* and *where* together, and situates these interactions within the context of wider retail sector processes, by highlighting relationships to external pressures (e.g., 'E-commerce'), arguably helping to better understand and effectively respond to them. Thus, the conceptual framework (chapter four), provides a new theoretical tool through which to better understand national retail (centre) systems, where integration and consideration of connections between retail centre geographies is key.

Substantive knowledge about retail (centre) geographies

In terms of substantive knowledge, this thesis studied the geographies of retail centres, generating new knowledge about the *who, what, where* as well as evidencing *how* they can be used as geographic data tools. Firstly, new knowledge about the spatiality of retail centres in the U.S. has been generated (chapters three and four), representing the first comprehensive set of retail centre boundaries for the U.S., joining up with similar efforts in the UK (e.g., Thurstain-Goodwin and Unwin, 2000; Pavlis et al., 2018; Macdonald et al., 2022). By extending such an analyses into a new international setting, it has been possible to review the spatial structure of retailing, and make direct comparisons to the UK. For instance, it is clear that U.S. retail centres are predominantly urbanised geographical phenomena, heavily concentrated in the most urbanised and heavily populated areas, with a greater density in the CBD or Downtown districts of major U.S. cities, as in the UK (Pavlis et al., 2018). Significant variation in the morphological forms of U.S. retail centres was also identified, with characteristically large sprawling centres in the CBD, and highly linear retail centres at the core of neighbourhoods and cities (similar to UK high streets), with more compact ones being typical of purpose-built developments like shopping centres and malls. Such insights contribute to existing literature on the spatial organisation of (retail) agglomerations, by highlighting that some of the structures that have existed historically such as "ribbons" and "centers" (Berry, 1963), remain pertinent in the contemporary urban environment. In relation to these, a significant contribution made in this thesis is that 'scale' remains a pertinent issue in U.S. retail centre geographies, where boundaries are of a much greater area than in the UK, thus, incorporation of land-use polygons and building footprints in chapter four was of significant utility, a notable implication for future investigation into the spatiality of urban phenomena in the U.S. This contrasts to much of the literature on the spatial organisation of retail agglomerations in the UK, where it has been sufficient to use point data only to capture the distribution of high streets and retail centres (Pavlis et al., 2018), although new approaches are starting to emerge that account for the growing scale of retail centres (Macdonald et al., 2022).

Secondly, a typological and non-hierarchical perspective on retail centres in the U.S. is presented for the first time (chapters three and four), which summarises *what* the characteristics of U.S. retail centres are. In particular, retail centres appear to exist within a two-tier system, with significant diversity in terms of their structural characteristics and location, as was the case in the original UK multidimensional typology (Dolega et al., 2021). For example, in both chapters three and four, a group characteristically found at the cores of urban areas and cities is seen, which are significantly different in characteristics to those found in other areas of the U.S., but similar to UK Regional and City centres (Pavlis et al., 2018). In contrast, many groups of retail centres were identified that are much more specific in terms of the characteristics, which have a much more specialist retail offering, but exhibit a more dispersed geographical distribution, similar to the "primary food and secondary comparison destinations" identified by Dolega et al. (2021). Thus, in terms of substantive contributions, it is clear that U.S. retail centres exhibit significant differences in terms of their characteristics and their geographic location, which have been better understood through consideration of them through a multidimensional perspective, as opposed to one based on differences in supply and demand.

Thirdly, in chapter four a national overview of retail centre patronage for the U.S. is provided, detailing *who* uses them and where they come from. Firstly, in terms of catchment behaviours, catchments exhibited noticeable differences across functions, with some of the largest retail centres exhibiting the largest catchments, noticeably apparent for the primary retail centres in the CBD of major American cities (e.g., Seattle), but less so with smaller retail centres or in polycentric cities (e.g., Los Angeles). This links directly to existing literature on patronage for retail centres; for example, in Lloyd and Cheshire (2017) the authors detailed the impacts of competing and nearby retail centres, which in some cases (e.g., Coleford) reduced the overall size of catchments, but in others (e.g., Monmouth) resulted in the incorporation or merging of the catchments of directly competing destinations. In addition, calibration of Huff model parameters (attractiveness, distance), generated significant new insights about the determinants of demand, a significant conceptual contribution highlighted earlier. Whilst much of the literature in the modelling of retail patronage has been concerned with formal definitions and conceptualisations of ‘attractiveness’ (Dolega et al., 2016; Gong et al., 2021; Jones et al., 2022), this thesis has shown that distance remains paramount to predicting the patronage of consumers to retail centres. This creates significant scope for how to measure, account for, and conceptualise ‘distance’, which has received some attention (e.g., Newing et al., 2015), but ought to be given greater consideration, especially when modelling patronage of retail centres. However, what remains clear is that a ‘one size fits all’ approach does not work for a national set of retail agglomerations, as in the UK (Dolega et al., 2016), and therefore retail centre functions and new forms of data need to be more explicitly accounted for when trying to understand patronage.

Finally, new knowledge has been generated about how retail centres have responded to external pressures, such as the COVID-19 pandemic, demonstrating in particular *how* retail centre geographies can be used as geographic data tools to understand the impacts of these pressures, building on existing efforts (e.g., Singleton et al., 2016; Comber et al., 2020; Trasberg and Cheshire, 2021). In chapter five, the recovery of retail centres following the initial shock of the pandemic was examined, through consideration of the role of *function*, region and *structural* characteristics of UK retail centres in determining their recovery trajectories over the last twelve months. This research contributed significantly to existing literature on the response of retail centres, which has thus far concentrated on earlier phases of the COVID-19 pandemic (e.g., Enoch et al., 2022; Frago, 2021), instead of examining how these responses have manifested in more recent phases. Chapter five identified significant heterogeneities in response, which documented the importance of modelling these individual recovery trajectories, and supported earlier findings that retail centre performance is inherently complex (Parker et al., 2017), exhibiting significant geographical differences and inequalities (Dolega and Lord, 2020; Singleton et al., 2016), as well as being related to the *functional* and *structural* characteristics of the retail centres themselves (Frago, 2021; Dolega and Lord, 2020; Enoch et al., 2022). Modelling of these relationships in chapter three revealed in particular that the *structural* (and catchment) characteristics of retail centres remain a greater determinant of recovery during recent months, with particularly negative consequences for retail centres in less deprived areas (as suggested by Dolega and Lord, 2020), lower resilience to online shopping (as suggested by Enoch et al., 2022) and higher proportions of leisure and fewer essential services, a trend observed much earlier in the pandemic (Ballantyne et al., 2021; Enoch et al., 2022). Thus, whilst significant uncertainty remains due to limitations of data used in chapter five (see Section 6.3), new insights about the recovery of retail centres have been obtained which contribute to existing narratives about the determinants of retail centre recovery and decline, whilst also highlighting the importance of multidimensional perspectives on retail centre performance, and providing an empirical basis upon which to continue to monitor it.

Policy implications

This thesis has made a number of methodological, substantive and theoretical contributions related to retail centre geographies, which have significant implications for the development of retail planning policy, but also in supporting efforts about how best to respond to some of the major challenges facing retail centres, as discussed in Section 2. Firstly, through examination of the geographies of U.S. retail centres, insights as to the existing provision of physical retailing space, which are often partial or incomplete, have now been obtained. These insights have great utility for policy and planning, through providing a more comprehensive overview of local, regional and national levels of retail provision, and particularly it's spatiality, characteristics and patronage behaviours. Thus, it is now more feasible to evaluate access to different types of retailing, identifying local areas that are underserved or under provisioned, ensuring equal access to vibrant and competitive local retail environments, through direct comparison of different retail centres, their catchments and performance over time. This is increasingly important given the pressures faced by the U.S. retail sector as it continues to traverse the 'retail apocalypse' (Helm et al., 2018), and emerges out of the COVID-19 pandemic, where it is expected that the vibrancy and vitality of local retail environments will have shifted dramatically.

Furthermore, whilst the geographies presented here are of a 'static' nature, constructed using data at a specific point of time, the analytical frameworks presented here offer significant potential to evaluate how local, regional, and national levels of retail provision are shifting over time, given replicable approaches are available (Dolega et al., 2021). In particular, this evidence provides a basis upon which to evaluate and monitor the performance of retail centres over time, building on existing efforts to conceptualise retail centre performance, as in related literature (e.g., Jones et al., 2022; Dolega and Lord, 2020; Philp et al., 2022), and based on the contributions of chapter five. Thus, for chapter four a GitHub repository has been constructed (see Ballantyne, 2022), which provides documented code through which to construct such geographies over time, providing insights as to how retail provision and the economic performance of retail centres is changing in the U.S., particularly in relation to the major challenges facing retail centres, and providing an important evidence base to support policy and planning through time.

Thirdly, the retail centre geographies presented in chapters three and four can provide an evidence base upon which to make effective decisions about the response of physical spaces of consumption to external pressures. A large body of work has used retail centre definitions, typologies, and catchments to understand how retail centres are responding to specific pressures, for example in Singleton et al. (2016), where the resilience of retail centres to 'E-commerce' was quantified through consideration of the interactions between supply and demand, and in Trasberg and Cheshire (2021), where retail centre boundaries and the accompanying multidimensional typology were used to highlight the differential response of consumption spaces to the early phases of the COVID-19 pandemic. Thus, coupling of retail centre geographies with ancillary datasets has helped to explore and quantify how some of these external pressures are modifying the retail (centre) system; retail centre geographies can be used to identify how and where effect responses are needed (Lloyd and Cheshire, 2017). Given the increasing role of experiential retailing in the 21st century (Grimsey, 2018), where retail centres and high streets are becoming places to do and experience rather than buy and purchase, such quantifications and insights are of great need, especially given the role of these spaces of consumption in the desirability of cities and urban areas (Glaeser et al., 2001). These insights can support development of legislation and the design of cities, to prioritise 'sense of place' and the quality of the retail offer (Baker and Wood, 2010), resulting in significant enhancements to the liveability and desirability of cities.

There has never been a greater need for such insights to support effective decision making, given the apparent and continued impacts of the COVID-19 pandemic on retail centres. Throughout this thesis, significant evidence has been documented that the pandemic has had significant impacts on retail centres, with particularly dramatic short-term consequences (chapter three), but also evidence of longer-term inequalities in recovery and decline towards the latter phases of the pandemic (chapter five). Such insights have significant implications for helping to plan and prepare these localities for similar shocks in the future (Enoch et al., 2022), through greater understanding of changes to consumer behaviour. Furthermore, given the complexity of retail centre responses (chapter three and five), it is vital that policies seeking to maintain or improve the vitality and viability of these spaces, adopt a multidimensional and holistic approach to doing so, beyond regional or *functional* strategies. However, without understanding the geographies of these retail centres first, in particular *where* they are located, *what* characteristics they have and *who* uses them, such interventions would not be possible, as the evidence base is not present to support them.

6.3. Study Limitations

It is important to consider the limitations of this study, to ensure interpretation of results and arising implications is robust. Whilst aims one and two explore the geographies of retail centres in the U.S., it is important to reiterate that they do not represent the definitive set of U.S. retail centre geographies. An empirical decision was made to exclude a significant proportion of retail centres delineated in both chapters, which negatively biases the smallest and most ruralised retail centres that are functionally very different, generating uncertainty about the conceptual framing of the retail centre system (see Figure 9). However, these decisions were necessary to capture the key geographies of retail centres in the U.S., in particular reducing the amount of 'noise' in chapter three, and limiting the computational resources needed to calibrate Huff models in chapter four. Secondly, the validation of retail centre geographies in chapter three and four was limited in that they did not engage with external stakeholders, practitioners, or town centre managers in the U.S. to validate them, unlike other examples where this has been commonplace (e.g., Macdonald et al., 2022; Dolega et al., 2021), to ensure the findings are representative of the U.S. retail system and of specific areas. However, given the nature of this PhD project, where constrained timescales and resources represented a significant challenge, it was not possible to carry out such an exercise. To resolve this limitation, the retail centre boundaries were validated against other relevant datasets (see Section 3.4), although there is generally a lack of suitable retail (or shopping) centre definitions in the U.S. Furthermore, expert knowledge from one of the co-authors of chapter three was utilised to validate our efforts, and ensure that the retail centre definitions and typology, as well as the approaches for deriving such outcomes were robust and fit-for-purpose in the U.S (see Sections 3.4 and 3.5). Finally, it is important to consider that these retail centre geographies are represented as something that can be solely defined in an empirical manner, and whilst much of the related research on this topic is empirical in nature, qualitative understandings about retail environments would arguably be of great utility, in verifying and validating the geographies presented, but also in suggesting additional considerations (i.e., input variables), based on customer or stakeholder experiences of the *who, what* and *where* of retail centres.

A second area of limitation relates to chapter five, in particular the approach to unpacking the recovery of retail centres based on exploratory analysis and a simple modelling framework. Firstly, it is difficult to ascertain for certain the exact nature of recovery, given the limitations of the mobility data as discussed below, so as a result are only able to make comparisons about recovery; considering how certain retail centres are recovering or declining relative to others. Secondly, the modelling framework was relatively simple in its consideration only of *function*, region and *structural* characteristics as explanatory variables, excluding well known impacts of seasonality, weather and holiday periods (Lyu et al., 2022; Rose and Dolega, 2022), non-hierarchical functions (Dolega et al., 2021) and the cost-of-living crisis (Wood, 2019). However, perhaps the most significant limitation relates to the Geolytix mobility data which has resulted in significant uncertainty as to the nature of recovery (Shi et al., 2022). In addition to well-known limitations with mobility data, the Geolytix mobility data is temporally unstable as a result of dramatic decreases in the sample of devices used to generate it. This has significantly reduced the ability to use it as a proxy economic performance measure. However, significant insights about working with unstable mobility data have been contributed, as following representativeness and stability analysis between different regions and *functions*, and use of the data as snapshots rather than a complete time-series, interesting and useful comparisons about recovery trajectories can be generated, which can have greater utility than other proxy measures like vacancy rates (Dolega and Lord, 2020) and footfall (Philp et al., 2022), which typically have constrained temporal availability or bias specific regions or locations. This creates significant potential also in the monitoring of the economic performance of retail centres in the U.S. too, as the SafeGraph data used to construct retail centre boundaries and typologies (chapters three, four) does not contain accurate information on whether individual units are vacant or not.

6.4. Future Research

The who, what and where of retail centre geographies

Chapters three and four have provided significant scope to explore the geographies of retail centres in other international settings, as this thesis has made such efforts empirically replicable. Chapter four presents a technique that can be used to delineate any urban agglomeration using available data, such as OpenStreetMap or SafeGraph as it expands its global coverage (SafeGraph Inc., 2021b). In terms of classification, such insights could be easily generated following a thorough review of relevant literature on the subject of study; for example, food retail environments (Shannon, 2016; Scharadin et al., 2022), before assembling comprehensive typologies and capturing specific 'niches' in the subject of study. Finally, if suitable data on observed patronage is available, whether from mobility datasets such as SafeGraph patterns or from other sources (e.g., loyalty cards) researchers can utilise the advancements made in Chapter four to estimate catchments for their subject of study, following customisation of the Huff model and calibration with suitable data to ensure observed behaviours are represented. These three potential avenues are now even more feasible as a GitHub repository is available for chapter four (see Ballantyne, 2022), providing researchers with code to delineate boundaries, develop typologies and generate catchments, and subsequently modify these approaches to suit their subject or location of study. For future studies into retail centre geographies, a conceptual framework of a national retail centre system has been established, which can be generalised to any retail environment, providing a conceptual and theoretical basis upon which to ground future research.

The how of retail centre geographies

This thesis has also provided another example of *how* retail centre geographies can be used as geographic data tools to understand the impacts of wider retail sector processes and external pressures on retail centres; through examination of their recovery following the COVID-19 pandemic, however this remains an area of great opportunity. Firstly, there is significant interest in further exploration of the response of retail centres to COVID-19, given availability of mobility (or other datasets) with greater temporal stability and availability. For example, modelling of retail centre trajectories prior to the pandemic until now would give clearer insights as to how exactly they have responded, especially through development of a sophisticated modelling framework that accounts for seasonality and other important influences, as discussed above. A particularly interesting research area would be in use of actual sales data to assess the economic performance of retail centres, as with store performance (e.g., Rose and Dolega, 2022), though such data is not yet available. Finally, looking forward, a 'Digital Twin' framework that capitalises on advancements in Big Data and modelling

(Goodchild, 2022; Shi et al., 2022), could be used to simulate how policy action can result in positive recovery and growth for retail centres as we emerge out of the pandemic. Such insights will be crucial as retail centres continue to face external threats, most recently in the form of the cost-of-living crisis. Whilst research is notably absent thus far, the impacts of the crisis are expected to be significant, with evidence that sales and footfall have already begun to decrease in recent months (ONS, 2022; Wright, 2022). Given the rising costs of energy and food, increasing taxes and falling wages in line with increasing inflation in the UK (Patrick and Pybus, 2022), the retail sector is expected to face some of the greatest challenges, with consumer confidence expected to fall and a subsequent wave of retail vacancies following it. Thus, whilst the pandemic is not over, the cost-of-living crisis represents another very significant 'shock' to the system, and further research could seek to unpack how retail centres are responding to it, taking advantage of insights gained in this thesis, notable around the value of retail centre geographies as geographic data tools through which to monitor these processes.

Evaluating and evidencing the utility of mobility data

Finally, building on the outcomes of chapters four and five, there is significant scope for more research about the use of mobility datasets, such as those from SafeGraph and Geolytix. As discussed, such datasets are subject to their own issues and biases, resulting in significant uncertainty about the findings produced, and their impacts on research, a fact that is too rarely underemphasised in published research, as argued by Trasberg and Cheshire (2021). However, once the reliability of these datasets is established, their utility generates endless research opportunities. As discussed throughout this thesis, mobility data is a useful source of information as it has a high spatio-temporal resolution, does not overly privilege specific geographic locations (e.g., anchor stores), and is collected passively. However, perhaps one of its greatest assets is that it can be used to assess the levels of activity in lots of different environments, unlike footfall which is generally restricted to those areas where sensors are available, typically high streets. For example, further research could use such data to understand the usage of green spaces and parks (Cui et al., 2022), which is a particularly pressing area of research following heightened use during the pandemic, and significant evidence linking them to positive mental health outcomes (Houlden et al., 2019). Thus, evaluating whether or not these trends of green space usage during the pandemic have stayed, and the implications of such trends on the future planning of green spaces and health of nearby residents is clearly a valuable area for future research. Thus, there is significant scope for mobility data both in the field of retail geography, and outside of it once researchers are able to effectively control for the limitations and biases created by such data. This thesis has arguably provided a useful starting point for doing so when working with unstable data, creating significant opportunities for the use of data like the Geolytix mobility dataset in a broad range of applications.

7. Bibliography

AbedRabbo, M., Hart, C. and Ellis–Chadwick, F. 2021. Do connected town-centre shopping journeys improve town-centre patronage? *International Journal of Retail & Distribution Management*, 49 (8): 1198-1212.

Astbury, G. and Thurstain-Goodwin, M. 2014. Measuring the impact of out-of-town retail development on town centre retail property in England and Wales. *Applied Spatial Analysis and Policy*, *7* (4): 301-316.

Baker, R. G., and Wood, S. 2010. Towards robust development of retail planning policy: Maintaining the viability and vitality of main street shopping precincts. *Geographical Research*, 48 (1): 65–74.

Baker, S. R., Farrokhnia, R. A., Meyer, S., Pagel, M., & Yannelis, C. 2020. How does household spending respond to an epidemic? Consumption during the 2020 COVID-19 pandemic. *The Review of Asset Pricing Studies*, *10* (4): 834–862.

Ballantyne, P. 2022. *USRetailCentres* (Version 1). [Source Code]. Available at: https://github.com/patrickballantyne/USRetailCentres.

Ballantyne, P. 2021. *RSRS* (Version 1). [Source Code]. Available at: https://github.com/patrickballantyne/RSRS.

Ballantyne, P., Singleton, A., Dolega, L., and Credit, K. 2022a. A framework for delineating the scale, extent and characteristics of American retail centre agglomerations. *Environment and Planning B: Urban Analytics and City Science*, 49 (3): 1112–28.

Ballantyne, P., Singleton, A., Dolega, L., and Macdonald, J. 2022b. Integrating the who, what and where of U.S. retail center geographies. *Annals of the American Association of Geographers.* https://doi.org/10.1080/24694452.2022.2098087.

Ballantyne, P., Singleton, A., Dolega, L. 2021. A Regional Exploration of Retail Visits During the COVID-19 Pandemic. *Regional Studies, Regional Science. 8* (1): 366-370.

Basker, E. 2007. The causes and consequences of Wal-Mart's growth. *Journal of Economic Perspectives*, *21* (3): 177-198.

Basu, A. and Muylle, S. 2003. Online support for commerce processes by web retailers. *Decision Support Systems*, *34* (4): 379-395.

Berman B. and Evans J.R. 2013. *Retail Management: A Strategic Approach*. 12th Edition, Pearson, Harlow, England.

Berry, B.J. 1963. *Commercial Structure and Commercial Blight: Retail Patterns and Processes in the City of Chicago*.

Berry, T., Newing, A., Davies, D. and Branch, K. 2016. Using workplace population statistics to understand retail store performance. *The International Review of Retail, Distribution and Consumer Research*, *26* (4): 375-395.

Birkin, M. and Heppenstall, A. 2011. Extending spatial interaction models with agents for understanding relationships in a dynamic retail market. *Urban studies research*, *2011:*

Birkin, M., Clarke, G. and Clarke, M. 2010. Refining and Operationalizing Entropy-Maximizing Models for Business Applications. *Geographical Analysis*, *42* (4): 422-445.

Boerschinger, O., Pansch, J. and Lupini, M. 2017. *Surviving the retail apocalypse: Charting the future of retail in Ramsey*. Resilient Communities Project, University of Minnesota. Available at: https://conservancy.umn.edu/handle/11299/193397 (Accessed 07/12/2022).

Bonaccorsi, G., Pierri, F., Cinelli, M., Flori, A., Galeazzi, A., Porcelli, F., Schmidt, A.L., Valensise, C.M., Scala, A., Quattrociocchi, W. and Pammolli, F. 2020. Economic and social consequences of human mobility restrictions under COVID-19. *Proceedings of the National Academy of Sciences*, *117* (27): 15530-15535.

Bonet-Morón, J., Ricciulli-Marín, D., Pérez-Valbuena, G. J., Galvis-Aponte, L. A., Haddad, E. A., Araújo, I. F. amd Perobelli, F. S. 2020. Regional economic impact of COVID-19 in Colombia: An input–output approach. *Regional Science Policy & Practice*, *12* (6): 1123–1150.
Borchert, J. G. 1998. Spatial dynamics of retail structure and the venerable retail hierarchy. *GeoJournal*, 45 (4): 327–36.
Brand Finance. 2022. The annual report on the most valuable and strongest luxury & premium brands Ranking. Available at: https://brandirectory.com/rankings/luxury-and-premium/table (Accessed 07/12/2022).
Braun, S. 2015. *The history of retail: a timeline.* Available at: https://www.lightspeedhq.co.uk/blog/the-history-of-retail-a-timeline/ (Accessed 07/12/2022).
Brown, S. 1992. *Retail location: a micro-scale perspective*. Ashgate Publishing Limited, Aldershot, England.
Burt, S., Sparks, L. and Teller, C. 2010. Retailing in the United Kingdom - a synopsis. *European Retail Research*: 173-194.
Calafiore, A., Murage, N., Nasuto, A., Rowe, F. 2022. *Deriving Spatio-Temporal Geographies of Human Mobility from GPS traces.* Available at: https://livrepository.liverpool.ac.uk/3146964/ (Accessed 07/12/2022).
Campello, R.J., Moulavi, D. and Sander, J. 2013. Density-based clustering based on hierarchical density estimates. In Pei, J. (ed.), *Advances in Knowledge Discovery and Data Mining*, Gold Coast, Australia: 160-172.
Campello, R.J., Moulavi, D., Zimek, A. and Sander, J. 2015. Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *10* (1): 1-51
Casparis, J. 1969. Shopping center location and retail store mix in metropolitan areas. *Demography*, 6 (2): 125–131.
Castree, N., Leszczynski, A., Stallins, J.A., Schwanen, T. and Patel, Z. 2022. Reconstituting geography for the 21st century. *Environment and Planning F*, *1* (1): 3-6.
CDRC, 2021a. *Geolytix aggregated in-app location dataset.* Available at: https://data.cdrc.ac.uk/dataset/geolytix-aggregated-app-location-dataset (Accessed 07/12/2022).
CDRC, 2021b. *Retail Centre Indicators*. Available at: https://data.cdrc.ac.uk/dataset/retail-centre-indicators (Accessed 07/12/2022).
Chowdhury, S., Bappy, M.H., Chowdhury, S., Chowdhury, M.S. and Chowdhury, N.S. 2022. Omicron Variant (B. 1.1. 529) of SARS-CoV-2, A Worldwide Public Health Emergency! *European Journal of Clinical Medicine*, *3* (1): 5-9.
Christaller, W. 1933. *Die Zentralen Orte in Suddeutscland.* Translated by Baskin, C.W., *Central Places in Southern Germany*, 1966.
Clapp, J. M., Ross, S. L., and Zhou, T. 2019. Retail agglomeration and competition externalities: Evidence from openings and closings of multiline department stores in the U.S. *Journal of Business & Economic Statistics*, 37 (1): 81–96.
Clark, W.A.V. 1967. The spatial structure of retail functions in a New Zealand city. *New Zealand Geographer*, *23* (1): 23-33.
Coca-Stefaniak, A. 2013. *Successful town centres: Developing effective strategies. Available* at: https://gala.gre.ac.uk/id/eprint/14927/ (Accessed 07/12/2022).
Comber, S., Arribas-Bel, D., Singleton, A., Dong, G. and Dolega, L. 2020. Building hierarchies of retail centers using Bayesian multilevel models. *Annals of the American Association of Geographers*, *110* (4): 1150-1173.
Cooke, R. 2018. Mary Portas: It was a question of how do I want to work as a woman? *The Guardian*, 21st October. Available at:

https://www.theguardian.com/lifeandstyle/2018/oct/21/mary-portas-interview-it-was-a-question-of-how-do-i-want-to-work-as-a-woman (Accessed 07/12/2022).
Credit, K. 2020. Neighbourhood inequity: Exploring the factors underlying racial and ethnic disparities in COVID-19 testing and infection rates using ZIP code data in Chicago and New York. *Regional Science Policy & Practice*, *12* (6): 1249-1271.
Crewe, L. 2000. Geographies of retailing and consumption. *Progress in human geography*, *24* (2): 275-290.
Cui, N., Malleson, N., Houlden, V. and Comber, A. 2022. Using social media data to understand the impact of the COVID-19 pandemic on urban green space use. *Urban Forestry & Urban Greening*, *74*: 127677.

Davies, A., Dolega, L. and Arribas-Bel, D. 2019. Buy online collect in-store: exploring grocery click & collect using a national case study. *International Journal of Retail & Distribution Management*, *47* (3): 278-291.
De Smith, M.J., Goodchild, M.F. and Longley, P. 2007. *Geospatial analysis: a comprehensive guide to principles, techniques and software tools*. Matador, Leicester, UK.
DeLisle, J.R. 2005. The Evolution of shopping center research: A 12-year retrospective. *Journal of Shopping Center Research*, *12* (2): 1-83.
Dennis, C., Marsland, D. and Cockett, T. 2002. Central place practice: shopping centre attractiveness measures, hinterland boundaries and the UK retail hierarchy. *Journal of Retailing and Consumer Services*, *9* (4): 185-199.
Department for Business Innovation and Skills. 2014. *BIS Research Paper No. 188: Policy Implications of Recent Trends in the High-Street/Retail Sector.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/389385/bis-14-1124_Research_paper_188_Policy_implications_of_recent_trends_in_the_high_street-_retail_sector.pdf (Accessed 07/12/2022).
Dixon, T. and Marston, A. 2002. UK retail real estate and the effects of online shopping. *Journal of Urban Technology*, *9* (3): 19-47.
Doherty, N.F. and Ellis-Chadwick, F. 2010. Internet retailing: the past, the present and the future. *International Journal of Retail & Distribution Management*, 38 (11/12): 943-965.
Dolega, L. and Celińska-Janowicz, D. 2015. Retail resilience: A theoretical framework for understanding town centre dynamics. *Studia Regionalne I Lokalne*, *60* (2): 8-31.
Dolega, L. and Lord, A. 2020. Exploring the geography of retail success and decline: A case study of the Liverpool City Region. *Cities*, *96*: 102456.
Dolega, L., Pavlis, M. and Singleton, A. 2016. Estimating attractiveness, hierarchy and catchment area extents for a national set of retail centre agglomerations. *Journal of Retailing and Consumer Services*, *28*: 78-90.
Dolega, L., Reynolds, J., Singleton, A. and Pavlis, M. 2021. Beyond retail: New ways of classifying UK shopping and consumption spaces. *Environment and Planning B: Urban Analytics and City Science*, *48* (1): 132-150.
Dramowicz, E. 2005. Retail trade area analysis using the Huff model. *Directions Magazine*, July 3rd. Available at: https://www.directionsmag.com/article/3207 (Accessed 07/12/2022).
Duong, B.V., Larpruenrudee, P., Fang, T., Hossain, S.I., Saha, S.C., Gu, Y. and Islam, M.S. 2022. Is the SARS CoV-2 Omicron Variant Deadlier and More Transmissible Than Delta Variant? *International Journal of Environmental Research and Public Health*, *19* (8): 4586.
Ellison, G., Glaeser, E.L. and Kerr, W.R. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, *100* (3): 1195-1213.

Elwood, S., Goodchild, M.F. and Sui, D.Z. 2012. Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of the association of American geographers*, *102* (3): 571-590.
Enoch, M., Monsuur, F., Palaiologou, G., Quddus, M.A., Ellis-Chadwick, F., Morton, C. and Rayner, R. 2022. When COVID-19 came to town: Measuring the impact of the coronavirus pandemic on footfall on six high streets in England. *Environment and Planning B: Urban Analytics and City Science*, *49* (3): 1091-1111.
Environmental Protection Agency (EPA). 2021. *Smart location database*. Available at: https://www.epa.gov/smartgrowth/smart-location-mapping#SLD (Accessed 07/12/2022).
Fetzer, T. 2020. Subsidizing the spread of Covid19: Evidence from the UK's Eat-Out To-Help-Out scheme. *The Economic Journal*, *132* (643): 1200-1217.
Fleming, D. 2009. *City of rhetoric: Revitalizing the public sphere in metropolitan America*. State University of New York Press, Albany, USA.

Fotheringham, A. S. 1983. A new set of spatial-interaction models: The theory of competing destinations. *Environment and Planning A: Economy and Space*, 15 (1): 15–36.
Frago, L. 2021. Impact of COVID-19 Pandemic on Retail Structure in Barcelona: From Tourism-Phobia to the Desertification of City Center. *Sustainability*, *13* (15): 8215.
Frances-Devine, B., Harari, D., Keep, M., Bolton, P., Harker, R., Barton, C. and Wilson, W. 2022. *Research Briefing: Rising cost of living in the UK.* Available at: https://commonslibrary.parliament.uk/research-briefings/cbp-9428/ (Accessed 07/12/2022).
Franklin, R.S. 2021. Geographical analysis at midlife. *Geographical Analysis*, *53* (1): 47-60.
Gao, S., Rao, J., Kang, Y., Liang, Y., Kruse, J., Dopfer, D., Sethi, A.K., Reyes, J.F., Yandell, B.S. and Patz, J.A. 2020. Association of mobile phone location data indications of travel and stay-at-home mandates with COVID-19 infection rates in the US. *JAMA Network Open*, *3* (9): 2020485.
Geiger, S. 2007. Exploring night-time grocery shopping behaviour. *Journal of Retailing and Consumer Services*, *14* (1): 24-34.
Geolytix. 2020. *Geolytix UK: Retail Place Boundaries*. Available at: https://geolytix.com/blog/uk-retail-places/ (Accessed 07/12/2022).
Gibbs, H., Waterlow, N. R., Cheshire, J., Danon, L., Liu, Y., Grundy, C., Kucharski, A. J., LSHTM CMMID COVID-19 Working Group and Eggo, R. M. 2021. Population disruption: estimating changes in population distribution in the UK during the COVID-19 pandemic. *medRxiv*. [Preprint]. June 24, 2021. Available from: https://doi.org/10.1101/2021.06.22.21259336.
Giraud, T., Cura, R. and Viry, M. 2020. *osrm: interface between R and the Open Street Map-based routing service (OSRM).* R package. Available at: https://cran.r-project.org/web/packages/osrm/index.html (Accessed 07/12/2022).
Glaeser, E.L., Kolko, J. and Saiz, A. 2001. Consumer city. *Journal of economic geography*, *1* (1): 27-50.
Gong, S., Cartlidge, J., Bai, R., Yue, Y., Li, Q., and Qiu, Q. 2021. Geographical and temporal huff model calibration using taxi trajectory data. *GeoInformatica*, 25 (3): 485–512.
Goodchild, M.F. 2022. Elements of an infrastructure for big urban data. *Urban Informatics*, *1* (1): 1-5.
Grabowski, F., Kochańczyk, M. and Lipniacki, T. 2022. The spread of SARS-CoV-2 variant Omicron with a doubling time of 2.0–3.3 days can be explained by immune evasion. *Viruses*, *14* (2): 294.
Grail, J. 2015. Strengthening the BID. *Property Journal*: 22-23.
Grewala, D., Roggeveena, A.L. and Nordfältb, J. 2017. The future of retailing. *Journal of retailing*, *93* (1): 1-6.

Grimsey, B. 2018. *The Grimsey Review 2: "It's time to reshape our town centres".* Available at: http://www.vanishinghighstreet.com/wp-content/uploads/2018/07/GrimseyReview2.pdf (Accessed 07/12/2022).
Grimsey, B., 2013. *The Grimsey Review: An Alternative Future For The High Street.* Available at: http://www.vanishinghighstreet.com/wp-content/uploads/2016/03/GrimseyReview04.092.pdf (Accessed 07/12/2022).
Guy, C. 2007. Planning for retail development. Routledge, London, UK.
Guy, C.M. 1998. Classifications of retail stores and shopping centres: some methodological issues. *GeoJournal*, *45* (4): 255-264.
Haklay, M., O'Sullivan, D., Thurstain-Goodwin, M. and Schelhorn, T. 2001. "So go downtown": simulating pedestrian movement in town centres. *Environment and planning B: Urban Analytics and City Science*, *28* (3): 343-359.
Halsall, S. 2001. *Identifying the Real Catchments of UK Shopping Locations.* Available at: https://www.geos.ed.ac.uk/~gisteac/proceedingsonline/AGI2001/TRACK%201/(3)From%20CRM%20to%20decision%20support/T1.4.pdf (Accessed 07/12/2022).
Han, Z., Cui, C., Miao, C., Wang, H. and Chen, X. 2019. Identifying spatial patterns of retail stores in road network structure. *Sustainability, 11* (17): 4539.
Harris, J.E. 2022. Mobility was a Significant Determinant of Reported COVID-19 Incidence During the Omicron Surge in the Most Populous U.S. Counties. *medRxiv.*
Helm, S., Kim, S.H. and Van Riper, S. 2020. Navigating the 'retail apocalypse': A framework of consumer evaluations of the new retail landscape. *Journal of Retailing and Consumer Services*, *54*: 101683.
High Streets Task Force (HSTF). 2021. *Review of High Street Footfall in England*. Available at: https://www.highstreetstaskforce.org.uk/media/hr5jbfev/footfall-report-2021-exec-summary-final.pdf (Accessed 07/12/2022).
HM Government. 2018. *Policy Paper: Future High Street Fund. Available* at: https://www.gov.uk/government/publications/future-high-streets-fund/future-high-street-fund (Accessed 07/12/2022).
HM Government. 2019. *New Task Force to help revitalise streets and town centres.* Available at: https://www.gov.uk/government/news/new-task-force-to-help-revitalise-high-streets-and-town-centres (Accessed 07/12/2022).
HM Government. 2022. COVID-19 Response: Living with COVID-19. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1056229/COVID-19_Response_-_Living_with_COVID-19.pdf. (Accessed 07/12/2022)
Houlden, V., de Albuquerque, J.P., Weich, S. and Jarvis, S. 2019. A spatial analysis of proximate greenspace and mental wellbeing in London. *Applied Geography*, *109*: 102036.
Huang, X., Lu, J., Gao, S., Wang, S., Liu, Z., and Wei, H. 2022. Staying at home is a privilege: Evidence from fine-grained mobile phone location data in the United States during the COVID-19 pandemic. *Annals of the American Association of Geographers,* 112 (1): 286–20.
Huff, D.L. 1964. Defining and estimating a trading area. *Journal of marketing*, *28* (3): 34-38.
Hutton, G. 2021. *Retail sector in the UK.* Available at: https://researchbriefings.files.parliament.uk/documents/SN06186/SN06186.pdf (Accessed 07/12/2022).
International Council of Shopping Centers (ICSC). 2017. *U.S. Shopping Center Classification and Characteristics.* Available at: https://www.icsc.com/uploads/research/general/US_CENTER_CLASSIFICATION.pdf (Accessed 07/12/2022).
Isidore, C. 2017. Retail's toughest year: A record for store closings. *CNN Business,* December 27. Available at: https://money.cnn.com/2017/12/26/news/companies/retail-toughest-year-store-closings/index.html (Accessed 07/12/2022).

Johnston, R., Harris, R., Jones, K., Manley, D., Wang, W.W. and Wolf, L. 2019. Quantitative methods I: The world we have lost–or where we started from. *Progress in Human Geography*, *43* (6): 1133-1142.

Jones, C. and Livingstone, N. 2018. The 'online high street' or the high street online? The implications for the urban retail hierarchy. *The International Review of Retail, Distribution and Consumer Research*, *28* (1): 47-63.

Jones, S., Newing, A. and Orford, S. 2022. Understanding town centre performance in Wales: using GIS to develop a tool for benchmarking. *Applied Spatial Analysis and Policy*, *15* (3): 713-740.

Joseph, L. and Kuby, M. 2016. The location types of US retailers. *International Journal of Applied Geospatial Research*, *7* (4): 1-22.

Kang, Y., Gao, S., Liang, Y., Li, M., Rao, J. and Kruse, J. 2020. Multiscale dynamic human mobility flow dataset in the US during the COVID-19 epidemic. *Scientific data*, *7* (1): 1-13.

Kerr, W.R. and Kominers, S.D. 2015. Agglomerative forces and cluster shapes. *Review of Economics and Statistics*, *97* (4): 877-899.

Kitchin, R. 2014. Big Data, new epistemologies and paradigm shifts. *Big data & society*, *1* (1).

Kohsaka, H. 1989. A spatial search-location model of retail centers. *Geographical analysis*, *21* (4): 338-349.

Lansley, G. and Longley, P. 2017. Opportunities for retail data and their geographic integration in social science, In Schintler, L.A. and Chen, Z (ed.) *Big Data for Regional Science*, Routledge, London, UK: 11-21.

Larsson, J. P., and Öner, O. 2014. Location and co-location in retail: A probabilistic approach using geo-coded data for metropolitan retail markets. *The Annals of Regional Science*, 52 (2): 385–408.

Latham, A., McCormack, D., McNamara, K. and McNeill, D. 2008. *Key Concepts in Urban Geography.* Sage Publishing, London, UK.

Liang, Y., Gao, S., Cai, Y., Foutz, N.Z. and Wu, L. 2020. Calibrating the dynamic Huff model for business analysis using location big data. *Transactions in GIS*, *24* (3): 681-703.

Lloyd, A. and Cheshire, J. 2017. Deriving retail centre locations and catchments from geo-tagged Twitter data. *Computers, Environment and Urban Systems*, *61*: 108-118.

Lüer-Villagra, A., Marianov, V., Eiselt, H.A. and Méndez-Vogel, G. 2022. The leader multipurpose shopping location problem. *European Journal of Operational Research*, *302* (2): 470-481.

Luxen, D. and Vetter, C. 2011. Real-time routing with OpenStreetMap data. In Agrawal, D et al., (ed.) *Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems*: 513-516.

Lyu, F., Wang, S., Han, S.Y., Catlett, C. and Wang, S. 2022. An integrated cyberGIS and machine learning framework for fine-scale prediction of Urban Heat Island using satellite remote sensing and urban sensor network data. *Urban Informatics*, *1* (1): 1-15.

Macdonald, J.L., Dolega, L. and Singleton, A. 2022. An open source delineation and hierarchical classification of UK retail agglomerations. *Scientific Data*, *9* (1): 1-12.

Martin, R. 2012. Regional economic resilience, hysteresis and recessionary shocks. *Journal of Economic Geography*, 12 (1): 1–32.

McArthur, E., Weaven, S. and Dant, R. 2016. The evolution of retailing: A meta review of the literature. *Journal of Macromarketing*, *36* (3): 272-286.

McCann, B.T., and Folta, T.B. 2008. Location matters: Where we have been and where we might go in agglomeration research. *Journal of Management*, 34 (3): 532–565.

McInnes, L. and Healy, J. 2017. Accelerated hierarchical density based clustering. In Gottumukkala et al., (ed.) *17th IEEE International Conference on Data Mining Workshops*, New Orleans, USA: 33-42.

McKee, L. 2014. *Retail Change: A Look at the Past, Present, and Future*. Masters thesis, North Dakota State University of Agriculture and Applied Science, Fargo, North Dakota. Available at: https://library.ndsu.edu/ir/bitstream/handle/10365/24753/Retail%20Change%20-%20A%20Look%20at%20the%20Past,%20Present,%20and%20Future.pdf?sequence=1 (Accessed 07/12/2022).
McMillen, D.P. 2003. Employment subcenters in Chicago: Past, present, and future. *Economic Perspectives*, 27 (2): 2–14.
Micu, C.B. 2019, September. Types of Shopping Centers—Literature Review. In Fotea, S.L., et al. (ed.) *Griffiths School of Management and IT Annual Conference on Business, Entrepreneurship and Ethics: Challenges and Opportunities to Develop Organisations Through Creativity, Technology and Ethics,* Oradea, Romania: 163-177.
Miller, H.J. and Goodchild, M.F. 2015. Data-driven geography. *GeoJournal*, *80* (4): 449-461.
Millington, S., Ntounis, N., Parker, C. and Quin, S. 2015. *Multifunctional Centres: a sustainable role for town and city centres*. Available at: https://www.highstreetstaskforce.org.uk/resources/details/?id=9f275023-d72f-488a-a61c-fc7a9b34ad4c (Accessed 07/12/2022).
Mu, L. 2008. A shape-based buffering method. *Environment and Planning B: Planning and Design*, *35* (3): 399-412.
Mumford, C., Parker, C., Ntounis, N. and Dargan, E. 2021. Footfall signatures and volumes: Towards a classification of UK centres. *Environment and Planning B: Urban Analytics and City Science*, *48* (6): 1495-1510.
Mumford, C.L., Parker, C.R., Ntounis, N. and Dargan, E. 2017. *A clustering study to verify four distinct monthly footfall signatures: A classification for UK retail centres*. Available at: https://v1.placemanagement.org/media/99861/monthlysignatureclusteringv45withlondon.pdf (Accessed 07/12/2022).
Murad, A.A. 2005. Using geographical information systems for exploring demand on retail centers at Jeddah city. In Zhao, H et al., (ed.) *Proceedings of ICSSSM'05. 2005 International Conference on Services Systems and Services Management,* Chongqing, China: 1523-1527.
Murphy, R.E. and Vance Jr, J.E. 1954. Delimiting the CBD. *Economic Geography*, *30* (3): 189-222.
National Retail Federation. 2019. Top 100 Retailers for 2019. Available at: https://nrf.com/resources/top-retailers/top-100-retailers/top-100-retailers-2019
Newing, A., Clarke, G.P. and Clarke, M. 2015. Developing and applying a disaggregated retail location model with extended retail demand estimations. *Geographical Analysis*, *47* (3): 219-239.
Newson, B. 2009. *Recession in the EU: its impact on retail trade.* Available at: https://edz.bib.uni-mannheim.de/www-edz/pdf/statinf/09/KS-SF-09-088-EN.pdf (Accessed 07/12/2022).
Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M. and Agha, R. 2020. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International journal of surgery*, *78*: 185-193.
Nong, Y., Zhou, S., Liu, L., Li, Q., Peng, Y. and Hao, X. 2019. Structural cities: delimiting retailing center boundaries and their hierarchical characteristics in urban China based on GPS-enabled taxi data. *Journal of Planning Education and Research*, *39* (3): 345-359.
Ntounis, N., Saga, R.S., Lorono-Leturiondo, M., Hindmarch, T., Parker, C. 2020. The Time to Act is Now: A Framework for Post-COVID 19 Recovery for our Towns and Cities. Available at: http://blog.placemanagement.org/2020/04/02/the-time-to-act-is-now-a-framework-for-post-covid-19-recovery-for-our-towns-and-cities/ (Accessed 07/12/2022).
O'Brien, L. 2020. h3jsr: Access Uber's H3 library. R package version 1.3.0. Available at: https://cran.r-project.org/web/packages/h3jsr/index.html (Accessed 07/12/2022).
O'Brien, L. and Harris, F. 1991. *Retailing: Shopping, Society, Space*. David Fulton Publishers, UK.

Office for National Statistics (ONS). 2022. *Retail sales, Great Britain: July 2022.* Available at: https://www.ons.gov.uk/businessindustryandtrade/retailindustry/bulletins/retailsales/july2022 (Accessed 07/12/2022).

Office of the Deputy Prime Minister (ODPM) and Centre for Advanced Spatial Analysis (CASA). 2002. *Producing Boundaries and Statistics for Town Centres: London Pilot Study Technical Report*. Available at: http://www.casa.ucl.ac.uk/towncentres/cd/Technical.pdf (Accessed 07/12/2022).

O'Kelly, M. and Bryan, D. 1996. Agricultural location theory: von Thunen's contribution to economic geography. *Progress in Human geography*, *20* (4): 457-475.

Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., De Nadai, M., Letouzé, E., Salah, A.A., Benjamins, R., Cattuto, C. and Colizza, V. 2020. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Science advances*, *6* (23).

Ooms, J. 2021. Magick: Advanced Graphics and Image Processing in R. R package version 2.7.3. Available at: https://cran.r-project.org/web/packages/magick/index.html (Accessed 07/12/2022).

Ordnance Survey (OS). 2019. *OS and ONS release report on the geography of Britain's high streets.* Available at: https://www.ordnancesurvey.co.uk/newsroom/blog/os-and-ons-release-report-on-the-geography-of-britains-high-streets (Accessed 07/12/2022).

Pan, S., Giannikas, V., Han, Y., Grover-Silva, E. and Qiao, B. 2017. Using customer-related data to enhance e-grocery home delivery. *Industrial Management & Data Systems*, *117* (9): 1917-1933.

Parker, C., Ntounis, N., Millington, S., Quin, S. and Castillo-Villar, F.R. 2017. Improving the vitality and viability of the UK High Street by 2020: Identifying priorities and a framework for action. *Journal of Place Management and Development*, *10* (4): 310-348.

Parr, J.B. 2017. Central place theory: An evaluation: Central Place Theory. *Review of urban and regional development studies*, *29* (3).

Parsons, E. 2020. *The ethics of geospatial – The four E's.* Available at: https://www.edparsons.com/2020/06/the-ethics-of-geospatial-the-four-es/ (Accessed 07/12/2022).

Patrick, R. and Pybus, K. 2022. Cost of living crisis: we cannot ignore the human cost of living in poverty. *bmj*, *377*.

Pavlis, M., Dolega, L. and Singleton, A. 2018. A modified DBSCAN clustering method to estimate retail center extent. *Geographical Analysis*, *50* (2): 141-161.

Philp, S., Dolega, L., Singleton, A. and Green, M. 2022. Archetypes of Footfall Context: Quantifying Temporal Variations in Retail Footfall in relation to Micro-Location Characteristics. *Applied Spatial Analysis and Policy*, *15* (1): 161-187.

Porta, S., Strano, E., Iacoviello, V., Messora, R., Latora, V., Cardillo, A., Wang, F. and Scellato, S. 2009. Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Urban Analytics and City Science*, *36* (3): 450-465.

Portas, M., 2011. *The portas review: An independent review into the future of our high streets*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/6292/2081646.pdf (Accessed 07/12/2022).

Pratt, M.D., Wright, J.A., Cockings, S. and Sterland, I. 2014. Delineating retail conurbations: A rules-based algorithmic approach. *Journal of Retailing and Consumer Services*, *21* (5): 667-675.

Prime Minister's Office. 2021. *Prime Minister confirms move to Plan B in England*. Available at: https://www.gov.uk/government/news/prime-minister-confirms-move-to-plan-b-in-england (Accessed 07/12/2022).

Pritzker, J.B. 2020. *Executive Order 2020–32 (COVID-19 Executive Order No. 30)*. Available at: www2.illinois.gov/Pages/Executive-Orders/ExecutiveOrder2020-32.aspx (Accessed 07/12/2022).
Proudfoot, M.J. 1937. City retail structure. *Economic Geography*, *13* (4): 425-428.
Rains, T. and Longley, P. 2021. The provenance of loyalty card data for urban and retail analytics. *Journal of Retailing and Consumer Services*, *63*: 102650.
Rao, F. 2020. Shopping centre morphologies in transition: towards a morphological typology of retail synergies. *Urban Design International*, *25* (4): 310-327.
Reilly, W.J. 1931. *The Law of Retail Gravitation.*
Reimers, V. and Clulow, V. 2009. Retail centres: it's time to make them convenient. *International Journal of Retail & Distribution Management*, *37* (7): 541-562.
Rey, S.J. 2015. Mathematical models in geography. In Wright, J.D. (ed.) *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*: 785-790.
Roggeveen, A.L. and Sethuraman, R. 2020. How the COVID-19 pandemic may change the world of retailing. *Journal of retailing*, *96* (2): 169.
Root, E. D. 2012. Moving neighborhoods and health research forward: Using geographic methods to examine the role of spatial scale in neighborhood effects on health. *Annals of the Association of American Geographers*, 102 (5): 986–95.
Rose, N. and Dolega, L. 2022. It's the Weather: Quantifying the Impact of Weather on Retail Sales. *Applied Spatial Analysis and Policy*, *15* (1): 189-214.
Rudavsky, R., Pollack, C.E. and Mehrotra, A. 2009. The geographic distribution, ownership, prices, and scope of practice at retail clinics. *Annals of internal medicine*, *151* (5): 315-320.
Sadahiro, Y. 2000. A PDF-based analysis of the spatial structure of retailing. *GeoJournal*, *52* (3): 237-252.
SafeGraph Inc. 2020a. *Places Manual.* Available at: https://docs.safegraph.com/docs/places-manual (Accessed 07/12/2022).
SafeGraph Inc. 2020b. *Weekly Patterns.* Available at: http://www.safegraph.com/covid-19-data-consortium (Accessed 07/12/2022).
SafeGraph Inc. 2020c. *Geometry data: The anchor of SafeGraph places.* Available at: https://www.safegraph.com/blog/geometry-the-anchor-of-safegraph-places (Accessed 07/12/2022).
SafeGraph Inc. 2021a. *SafeGraph's Data Sourcing Process*. Available at: https://www.safegraph.com/blog/safegraphs-data-sourcing-process (Accessed 07/12/2022).
SafeGraph Inc. 2021b. *Announcing SafeGraph global brands.* Available at: https://www.safegraph.com/blog/announcing-safegraph-global-brands (Accessed 07/12/2022).
SafeGraph Inc. 2022. *Site de-selection: Choosing the right stores to close comes down to data.* Available at: https://www.safegraph.com/blog/site-de-selection-choosing-the-right-stores-to-close-comes-down-to-data. (Accessed 07/12/2022).
Scharadin, B., M. Ver Ploeg, and C. Dicken. 2022. Geographic boundary definitions and the robustness of common food retail environment measures. *Annals of the American Association of Geographers*, 112 (5): 1403–23.
Segal, D.B. 1999. Retail trade area analysis: concepts and new approaches. *Journal of Database Marketing*, *6*: 267-278.
Sevtsuk, A. 2014. Location and agglomeration: The distribution of retail and food businesses in dense urban environments. *Journal of Planning Education and Research*, 34 (4): 374–93.
Shannon, J. 2016. Beyond the supermarket solution: Linking food deserts, neighborhood context, and everyday mobility. *Annals of the American Association of Geographers*, 106 (1): 186–202.

Shi, W., Batty, M., Goodchild, M. and Li, Q. 2022. The digital transformation of cities. *Urban Informatics*, *1* (1): 1-2.
Simbolon, R. and Riyanto, S. 2020. How Retail Survive Against Pandemic of COVID-19: An Insight from Optical Retailer. *International Journal of Current Science and Multidisciplinary Research*, *3*: 124-132.
Singleton, A. and Arribas-Bel, D. 2021. Geographic data science. *Geographical Analysis*, *53* (1): 61-75.
Singleton, A., Alexiou, A.,and Savani, R. 2020. Mapping the geodemographics of digital inequality in Great Britain: An integration of machine learning into small area estimation. *Computers, Environment and Urban Systems*, 82: 101486.
Singleton, A.D., Dolega, L., Riddlesden, D. and Longley, P.A. 2016. Measuring the spatial vulnerability of retail centres to online consumption through a framework of e-resilience. *Geoforum*, *69:* 5-18.
Singleton, A.D., Longley, P.A., Allen, R. and O'Brien, O. 2011. Estimating secondary school catchment areas and the spatial equity of access. *Computers, Environment and Urban Systems*, *35* (3): 241-249.
Sit, K.J., Ballantyne, E.E. and Gorst, J. 2022. Profiling shoppers' coping behaviours during a pandemic crisis: A regulatory focus perspective. *Journal of Retailing and Consumer Services*, *64*: 102811.
Sohn, J., Kim, T.J., and Hewings, G.J.D. 2003 Information technology and urban spatial structure: A comparative analysis of the Chicago and Seoul regions. *The Annals of Regional Science*, 37 (3): 447–462.
Spielman, S.E. and Singleton, A. 2015. Studying neighborhoods using uncertain data from the American community survey: a contextual approach. *Annals of the Association of American Geographers*, *105* (5): 1003-1025.
Struyf, A., Hubert, M. and Rousseeuw, P. 1997. Clustering in an object-oriented environment. *Journal of Statistical Software*, *1:* 1-30.
Teller, C. and Reutterer, T. 2008. The evolving concept of retail attractiveness: what makes retail agglomerations attractive when customers shop at them?. *Journal of Retailing and Consumer Services*, *15* (3): 127-143.
Tennekes, M., Nowosad, J., Gombin, J., Jeworutzski, S., Russell, K., Zijdeman, R., Clouse, J., Lovelace, R., and Muenchow, J. 2021. tmap: Thematic Maps. R package version 3.3-3. Available at: https://cran.r-project.org/web/packages/tmap/index.html (Accessed 07/12/2022).
The Magnificent Mile Association. 2015. *Vision 2025: A Policy Plan for North Michigan*. Available at: www.themagnificentmileassociation.com/assets/PDF/d9439ec65d/Vision-2025-July-2015-reduced.pdf (Accessed 07/12/2022).
Thompson, A. and Walker, J. 2005. Retail network planning—achieving competitive advantage through geographical analysis. *Journal of Targeting, Measurement and Analysis for Marketing*, *13* (3): 250-257.
Thompson, C.P. 2013. *Retail spending and store location during a recession: an analysis of changing consumer behaviour and interaction patterns*. PhD Thesis, University of Leeds, Leeds, UK. Available at: https://core.ac.uk/download/pdf/20482325.pdf (Accessed 07/12/2022).
Thurstain-Goodwin, M. and Unwin, D. 2000. Defining and delineating the central areas of towns for statistical monitoring using continuous surface representations. *Transactions in GIS*, *4* (4): 305-317.
Torres-Baron, V. 2018. *Is American retail at a historic tipping point?.* Available at: https://academicworks.cuny.edu/gj_etds/296/.

Torrisi, G. 2020. COVID: Everything is better when everything is worse? A comparative analysis of testing, death rate and life expectancy across Italian regions. *Regional Studies, Regional Science*, *7* (1): 463–475.
Trasberg, T. and Cheshire, J. 2021. Spatial and social disparities in the decline of activities during the COVID-19 lockdown in Greater London. *Urban Studies*, *0* (0): 211-236.
Tselios, V., Lambiri, D. and Dolega, L. 2018. Performance within a recession: The converging trajectories of retail centres in the UK. *Regional Science Policy & Practice*, *10* (4): 347-365.
U.S. Census Bureau. 2021. *American Community Survey data.* Available at: https://www.census.gov/programs-surveys/acs/data.html (Accessed 07/12/2022).
Uber. 2018. *H3: Uber's Hexagonal Hierarchical Spatial Index*. Available at: https://eng.uber.com/h3/ (Accessed 07/12/2022).
US Census Bureau. 2018. *LEHD origin destination employment statistics – Workplace area characteristics.* Available at: https://lehd.ces.census.gov/data/lodes/LODES7/il/wac/ (Accessed 07/12/2022).
Vom Hofe, R., and Bhatta, S.D. 2007. Method for identifying local and domestic industrial clusters using interregional commodity trade data. *Industrial Geographer* 4 (2): 1–27.
Von Thünen, J.H. 1826. *Der Isolierte Staaat in Beziehung auf Landtschaft und Nationalokonomie*.
Waddington, T.B., Clarke, G.P., Clarke, M. and Newing, A. 2018. Open all hours: Spatiotemporal fluctuations in UK grocery store sales and catchment area demand. *The International Review of Retail, Distribution and Consumer Research*, *28* (1): 1-26.
Wang, F., Chen, C., Xiu, C. and Zhang, P. 2014. Location analysis of retail stores in Changchun, China: A street centrality perspective. *Cities*, *41*: 54-63.
Wang, J., N. McDonald, A. L. Cochran, L. Oluyede, M. Wolfe, and L. Prunkl. 2021. Health care visits during the COVID-19 pandemic: A spatial and temporal analysis of mobile device data. *Health & Place*, 72: 102679.
Wang, Y., Wang, T., Tsou, M.H., Li, H., Jiang, W. and Guo, F. 2016. Mapping dynamic urban land use patterns with crowdsourced geo-tagged social media (Sina-Weibo) and commercial points of interest collections in Beijing, China. *Sustainability*, *8* (11): 120.
Weltevreden, J.W. 2007. Substitution or complementarity? How the Internet changes city centre shopping. *Journal of Retailing and consumer Services*, *14* (3): 192-207.
Widaningrum, D.L. 2015. A GIS–based approach for catchment area analysis of convenience store. *Procedia Computer Science*, *72*: 511-518.
Wieland, T. 2017. Market Area Analysis for Retail and Service Locations with MCI. *R Journal*, *9* (1).
Wikipedia. 2022. Discount Stores of the United States. Available at: https://en.wikipedia.org/wiki/Category:Discount_stores_of_the_United_States (Accessed 07/12/2022).
Williams, D.E. 2009. The evolution of e-tailing. *The International Review of Retail, Distribution and Consumer Research*, *19* (3): 219-249.
Wilson, A. G. 1969. The use of entropy maximising models, in the theory of trip distribution, mode split and route split. *Journal of Transport Economics and Policy*, 3 (1):108–26.
Wood, Z. 2019. Boxing Day sales dip blamed on poor weather and Black Friday. *The Guardian*, 26th December. Available at: https://www.theguardian.com/business/2019/dec/26/boxing-day-footfall-dip-blamed-on-poor-weather-and-black-friday (Accessed 07/12/2022).
Woodbury, C. 1928. The size of retail business districts in the Chicago metropolitan region. *Journal of Land & Public Utility Economics*, *4*: 85.
Wright, G. 2022. July UK footfall shows "north-south divide emerging" amid the burden of inflation. *Retail Gazette*, August 4th. Available at: https://www.retailgazette.co.uk/blog/2022/08/july-uk-footfall-shows-north-south-divide-emerging-amid-the-burden-of-inflation/ (Accessed 07/12/2022).

Wrigley, N. and Lowe, M. 2002. *Reading Retail.* Arnold Publishing, London, UK.
Wrigley, N., Lambiri, D., Astbury, G., Dolega, L., Hart, C., Reeves, C., Thurstain-Goodwin, M. and Wood, S. 2015. *British high streets: From crisis to recovery? A comprehensive review of the evidence.* Available at: https://eprints.soton.ac.uk/375492/ (Accessed 07/12/2022).
Yabe, T., P. S. C. Rao, and S. V. Ukkusuri. 2021. Regional differences in resilience of social and physical systems: Case study of Puerto Rico after Hurricane Maria. *Environment and Planning B: Urban Analytics and City Science,* 48 (5): 1042–57.
Yilmazkuday, H. 2020. Stay-at-home works to fight against COVID-19: International evidence from Google mobility data. *Journal of Human Behavior in the Social Environment*, *31* (1-4): 210-220.
Yoshimura, Y., Santi, P., Arias, J.M., Zheng, S. and Ratti, C. 2021. Spatial clustering: Influence of urban street networks on retail sales volumes. *Environment and Planning B: Urban Analytics and City Science*, *48* (7): 1926-1942.
Yrigoyen, C.C. and Otero, J.V. 1998. *Spatial interaction models applied to the design of retail trade areas.* Available at: https://www.econstor.eu/handle/10419/113417 (Accessed 07/12/2022).
Zhai, W., Liu, M., Fu, X., and Peng, Z.R. 2021. American inequality meets COVID-19: Uneven spread of the disease across communities. *Annals of the American Association of Geographers,* 111 (7): 2023–43.
Zhou, T., and Clapp, J.M. 2015. The location of new anchor stores within metropolitan areas. *Regional Science and Urban Economics,* 50: 87–107.

8. A Regional Exploration of Retail Visits during the COVID-19 Pandemic.

The content of this paper was published as a Special Issue in Regional Studies, Regional Science:

Ballantyne, P., Singleton, A., Dolega, L., 2021. A Regional Exploration of Retail Visits during the COVID-19 Pandemic. *Regional Studies, Regional Science. 8* (1), 366-370.

DOI: https://doi.org/10.1080/21681376.2021.1973548

Abstract

Despite evidence that the COVID-19 pandemic has precipitated significant regional (economic) inequalities, there is a substantial lack of regional insight into the impacts of COVID-19 on the retail sector. In this study, using data from SafeGraph, a regional approach is adopted to explore how visits to retail places changed during the early weeks of the COVID-19 pandemic in the Chicago Metropolitan area. In particular, this study highlights that retail visits exhibited interesting spatio-temporal and structural trends.

8.1. Background

A wealth of research is rapidly emerging that seeks to understand the interactions between COVID-19 and retail, specifically quantifying the impacts of reduced mobility (due to national lockdowns and restrictions) on the economic performance of the sector (Yilmazkuday, 2020; Baker et al., 2020). Many of these studies have used novel datasets (e.g. Google Mobility), identifying significant shifts in mobility and expenditure between different types of retail. However, despite evidence that the pandemic has precipitated significant regional (economic) inequalities between different sectors (Bonet-Morón et al., 2020), and that COVID-19 is an inherently regional issue (Torrissi, 2020); there is a substantial lack of (regional) insight into the impacts of COVID-19 on the retail sector.

Here regional approach is adopted, focusing on the Chicago MSA in the 'Mid-West' region of the USA. Change in total visits to retail 'places' during the early weeks of the COVID-19 pandemic is explored, before unpacking these trends further by considering how they relate to specific types of retail (e.g. convenience). This piece is important, providing both an insight into the response of a regional retail sector to COVID-19, whilst also demonstrating the utility of mobility datasets and spatial indexing systems like ‘H3’ (Uber, 2018), at conveying trends in mobility, whilst also preserving the security of store-level data.

8.2. Trends in Regional Retail Visits

Retail visits followed an uneven spatial distribution across the Chicago MSA (Figure 27). Irrespective of week, the vast number of visits were concentrated in and around the CBD of Chicago. Other significant but smaller concentrations were found in ‘satellite cities’ like Joliet, and in established shopping parks, such as the Woodfield and Fox Valley Malls. This is interesting as it generates the potential to unpack an approximate geography of retail space using ‘Big Data’ (Lloyd and Cheshire, 2017), rather than delineation of the retail locations themselves (Pavlis et al., 2018). Temporally, there was a clear trend of decreasing retail visits which aligned closely with the first ‘peak’ of the pandemic (Baker et al., 2020), suggesting that as the pandemic worsened, people began to alter their consumer behaviour, visiting stores less frequently. The most significant decrease in visits occurred in the week where the ‘Stay at Home’ order was issued (16/03), which permitted residents of Illinois to only leave their homes for ‘essential’ activities (Pritzker, 2020). Furthermore, an evident spatio-temporal pattern was that the suburban/rural parts of Chicago MSA appeared to experience greater contractions in visits when compared to the CBD, prompting a future research agenda to better understand why this might be.

To unpack further some of the trends identified in Figure 27, this study explores how these variations in visits related to different types of retail (Figure 28). Convenience retail (e.g., grocery stores) saw a substantial and sustained increase in the proportion of visits from 28 to 35% following the ‘Stay at Home’ order, likely a result of increased demand of ‘essential goods’ (e.g., groceries), characteristic of the early weeks of the pandemic (Nicola et al., 2020). Another interesting trend was the significant and sustained decline leisure-based retail visits (8%), a component of the retail sector (e.g., restaurants) that has faced some of the greatest impacts during the pandemic (Baker et al., 2020).

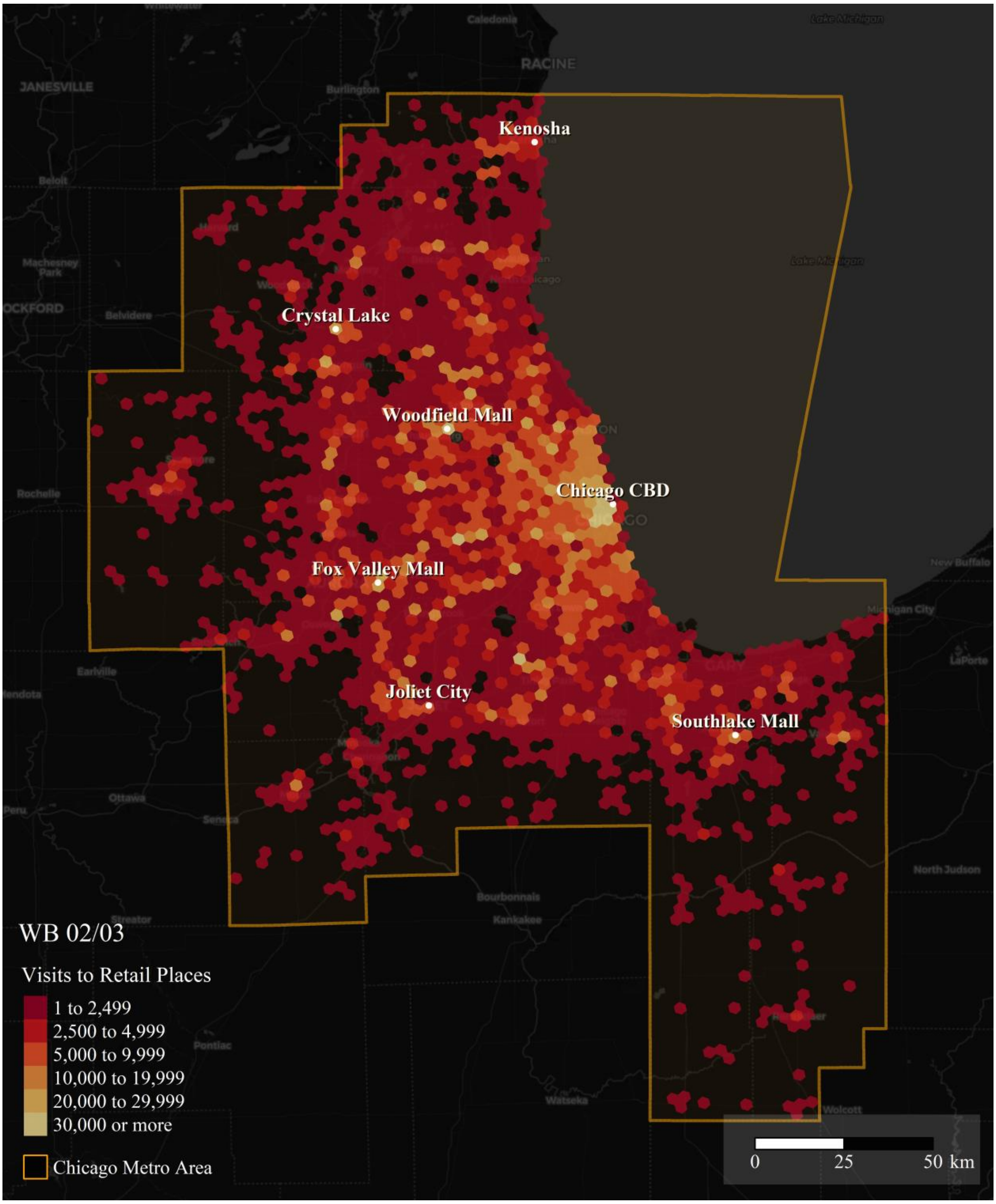

Figure 27. Weekly visits to retail places in the Chicago MSA, from the week beginning 2nd March to the 6th April. Each iteration represents one week of data (e.g. WB 02/03), and the general trend seen across the iterations is a decrease in retail visits throughout the MSA in these early weeks, with the most dramatic decrease coinciding with the issuing of the 'Stay at Home' order (WB 16/03).

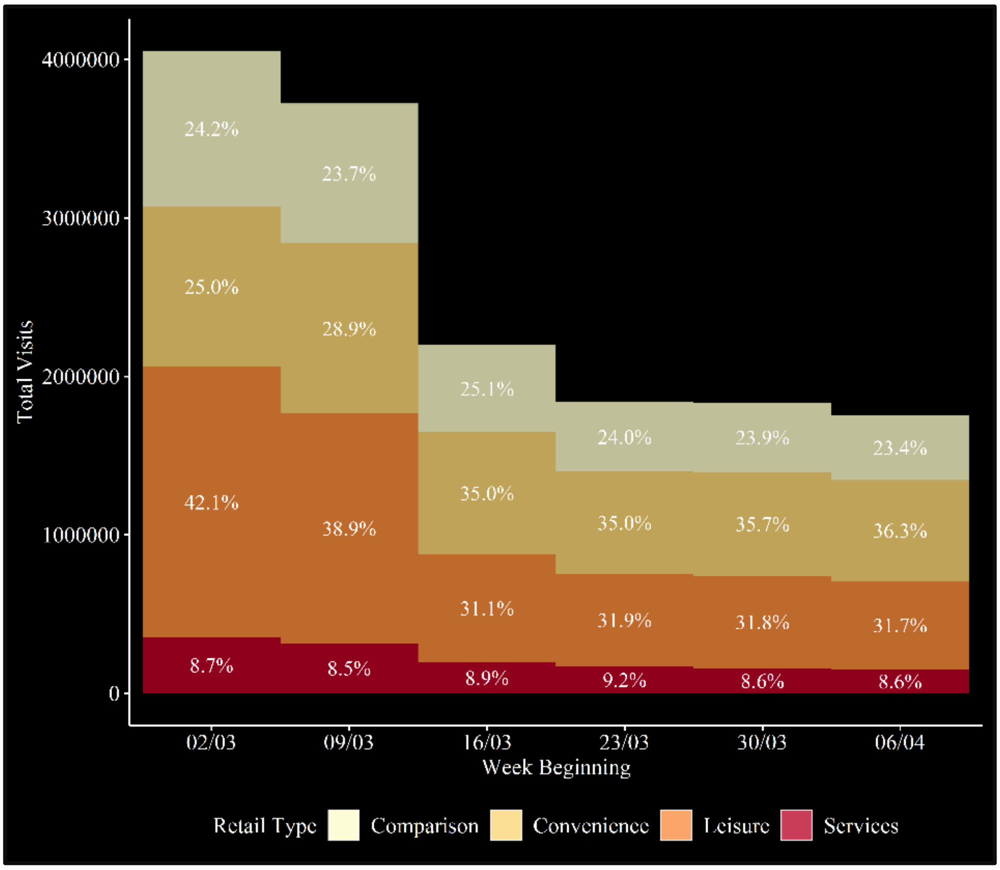

Figure 28. Weekly visits to retail places disaggregated by retail type. This figure highlights in particular the general increased uptake of 'essential' goods (e.g. convenience), following issuing of the 'Stay at Home' order (WB 16/03), and decreased usage of leisure-based retail (e.g. drinking places and restaurants). For information about the retail types and their aggregation see Table 13 and its accompanying caption.

From these findings, it is clear that further research is required to unpack these regional trends in retail visits, in particular seeking to quantify the role of two evident drivers in determining consumer mobility during the pandemic: geographical context (Figure 27) and retail offering (Figure 28). Modelling of retail visits in relation to both the offering (e.g. retail type, brands) and geographical context (e.g. urban vs rural, COVID-19 infections) of retail locations would likely yield significant insights as to the drivers of such trends. Although beyond the scope of this exploratory analysis, these insights could have significant merit in retail planning. For example, underperforming neighbourhoods and/or components of the retail sector could be targeted with post-pandemic economic recovery strategies, such as the 'Eat Out to Help Out' scheme that was introduced in the UK to support the hospitality and restaurant sector during the summer of 2020 (Fetzer, 2020).

8.3. Technical Details

Through the 'COVID-19' Data Consortium, SafeGraph have provided researchers with access to their datasets, including a register of 'core places' where consumers spend money or time (SafeGraph Inc, 2020a), and corresponding mobility data ('weekly patterns') collected from the GPS data of 45 million anonymised phone users (SafeGraph Inc, 2020b). From the 'core places' dataset, those places pertaining to retail were extracted, by matching the SafeGraph 'top categories' to one of four 'retail types', and removing all other 'core places' not assigned to a 'retail type', as seen in Table 13.

The 'retail places' were then joined with their corresponding 'weekly patterns', in particular data for a six-week period surrounding the first peak of the pandemic (week beginning 2nd March to 6th April). The places (and patterns) were then joined onto a hexagonal grid for the Chicago MSA, constructed using the 'h3jsr' R package (O'Brien, 2020), enabling visualisation of change in weekly visits across Chicago (Figure 27), using the tmap and magick R packages (Tennekes et al., 2021; Ooms, 2021). Figure 2 was constructed by calculating the proportion of total weekly visits occupied by each of the four retail types. The code used to produce these outputs can be found on the authors GitHub (see Ballantyne, 2021).

Table 13. Aggregation of 'core places' to retail types, enabling extraction of 'retail places' from the dataset, based loosely on aggregations in the LDC 'Retail Type and Vacancy' dataset (LDC, 2020). 'Comparison' relates to less frequently purchased, non-food retail (e.g., fashion, gifts), 'convenience' retail relates to frequently purchased 'essential goods' (e.g., food, drink, fuel). 'Leisure' retail incorporates all form of entertainment (e.g. drinking places, theatres), and 'service' retail covers all services/utilities offered in retail spaces.

Retail Type	SafeGraph Top Category (Examples
Comparison	Clothing Stores, Department Stores, Furniture Stores, Automobile Dealers, Electronics and Appliance Stores, Office Supplies, Stationery and Gift Stores.
Convenience	Grocery Stores, Gasoline Stations, Health and Personal Care Stores, Beer Wine & Liquor Stores, General Merchandise Stores, Speciality Food Stores.
Leisure	Drinking Places, Restaurants and Other Eating Places, Motion Picture and Video Industries, Gambling Industries, Traveller Accommodation.
Service	Automotive Repair and Maintenance, Personal Care Services, Insurance Carriers, Depository Credit Intermediation, Taxi and Limousine Services.

Acknowledgements

We would like to thank SafeGraph Inc. for permitting access to various SafeGraph datasets ('core places', 'weekly patterns') as part of their COVID-19 Data Consortium.

Appendices

I. A framework for delineating the scale extent and characteristics of American retail centre agglomerations

A. SafeGraph Datasets

The SafeGraph ‘core places’ dataset consists of approximately 6.3 million POIs for the US, containing information about the POIs, summarised below in Table 14. The SafeGraph ‘core places’ dataset is updated and re-released each month to reflect openings/closures of businesses (SafeGraph Inc., 2021a), providing what is likely the most comprehensive and up-to-date, open-access database of businesses in the US. It is compiled in a variety of ways; web crawling of major brands/businesses, extracting POI locations from public APIs and open web domains, and licensing of third-party data (SafeGraph Inc., 2021a). An intense cleaning process is used to remove and/or merge POI duplicates, and once the final POIs and their locations have been extracted, a machine-learning model and existing metadata are used to accurately assign a 6-digit NAICS (North American Industrial Classification System) code to each POI (SafeGraph Inc., 2021a). This NAICS code and accompanying label is used to derive the ‘top categories’ (using the first four digits, e.g., NAICS 4451 = Grocery Stores), and ‘subcategories’ (using the six digits, e.g. NAICS 445120 = Convenience Stores).

Table 14. Structure of the SafeGraph 'core places' dataset.

Variable	Description	Source
safegraph_place_id	Unique ID tied to each POI	Assigned uniquely post-cleaning (removal/merging of duplicates).
location_name	Name of the POI	Extracted from web-crawling, public APIs/open web domains and third-party data.
brands	Brand name if POI belongs to a larger brand	Extracted from web-crawling, public APIs/open web domains and third-party data.
top_category	Label associated with the first 4 digits of the NAICS code assigned to that POI	NAICS code assigned to each POI using a machine learning model and existing metadata. The top category is then assigned using the corresponding labels for the

		first 4 digits of that NAICS code.
sub_category	Label associated with the 6 digits of the NAICS code assigned to that POI	NAICS code assigned to each POI using a machine learning model and existing metadata. The sub category is then assigned using the corresponding labels for the 6 digits of that NAICS code.
latitude	Latitude coordinates of the POI	Defined by best knowledge of the POI location, aiming to identify the 'centre of the business', whilst also capturing the accurate locations within larger businesses (e.g. strip malls).
longitude	Longitude coordinates of the POI	Defined by best knowledge of the POI location, aiming to identify the 'centre of the business', whilst also capturing the accurate locations within larger businesses (e.g. strip malls).

To extract the ‘retail places’ from the ‘core places’ dataset, the SafeGraph ‘top categories’ were aggregated to the retail categories used in Pavlis et al. (2018), based on the LDC ‘Retail Type/Vacancy’ dataset (Table 15). As there were only 175 distinct ‘top categories’ in the SafeGraph ‘core places’ for Chicago MSA, the aggregation was completed manually by ascribing a ‘retail aggregation’ (Table 15) based on sensible overlaps between the SafeGraph ‘top categories and the corresponding POI categories and higher-level aggregation in the LDC dataset. Once every POI had been aggregated to one of the categories seen below in Table 15, the non-retail features were removed.

Table 15. Example aggregation of SafeGraph 'Top Categories' to identify retail types and non-retail places, and their relationship to the NAICS system.

Retail Aggregation (based on Pavlis et al. and LDC)	**Example SafeGraph 'Top Categories'**	**NAICS Codes**
Comparison	Automobile Dealers, Furniture Stores, Clothing Stores, Department Stores	4481, 4411, 4421, 4522
Convenience	Grocery Stores, Speciality Food Stores, Health and Personal Care Stores, Gasoline Stations	4451, 4452, 4471, 4461

Leisure	Spectator Sports, Amusement Parks & Arcades, Drinking Places, Restaurants and Other Eating Places	7112, 7131, 7224, 7225
Service	Depository Credit Intermediation, Insurance Carriers, Automotive Repair and Maintenance, Personal Care Services	5221, 5241, 8111, 8121
Non-Retail	Waste Treatment and Disposal, Technical and Trade Schools, Offices of Physicians, Individual and Family Services	5662, 6115, 6211, 6241

The ‘weekly patterns’ dataset provides foot traffic data for the ‘core places’. It is constructed by licencing aggregated and anonymised mobility GPS data from companies like Veraset, which obtain data from unnamed mobile applications in which users give permissions to share their location (SafeGraph Inc., 2021a). Following compilation of the raw GPS data from the various sources, SafeGraph anonymises it and processes it to aggregate place-level variables such as ‘total visits’ and ‘median dwell’ for its ‘core places’ (SafeGraph Inc., 2021a). A summary of what this looks like can be seen below in Table 16. The aggregation of GPS data to the ‘core places’ is completed by clustering the GPS data and then using a machine learning model to predict the best possible ‘place’ for each cluster of GPS data (SafeGraph Inc., 2021a).

The ‘weekly patterns’ dataset had two purposes in this paper. Firstly, it was used to derive retail centre variables for the creation of a typology in Section 3.5; the ‘weekly patterns’ for the ‘retail places’ in each retail centre were extracted for a series of weeks prior to the COVID-19 pandemic, and the most ‘typical’ week was used to derive retail centre variables for the typology (Table 16), as described below in section C. The ‘weekly patterns’ dataset was also used in Section 3.5; the total visits for each retail centre were calculated, before aggregating these by retail centre group, as identified in the previous section. This enabled weekly share (%) in total visits to retail centres to be calculated between the different groups of centres, enabling comparison of share change from September 2019 – September 2020.

Table 16. Structure of SafeGraph 'weekly patterns' dataset.

Variable	Description	Source
safegraph_place_id	Unique ID tied to each POI	Asssigned uniquely post-cleaning (removal/merging of duplicates)
location_name	Name of the POI	Extracted from web-crawling, public APIs/open web domains and third-party data
date_range_start	Start date and time for measurement week (yyyy-mm-ddT00:00-06:00)	Processing of raw GPS data from licensed third-party datasets
date_range_end	End date and time for measurement week (yyyy-mm-ddT00:00-06:00)	Processing of raw GPS data from licensed third-party datasets
raw_visit_counts	Number of visits to this POI during the date and time range	Processing of raw GPS data from licensed third-party datasets and aggregation to 'core places'
raw_visitor_counts	Number of unique visitors to this POI during the date and time range	Processing of raw GPS data from licensed third-party datasets and aggregation to 'core places'
distance_from_home	Median distance travelled by visitors in meters	Processing of raw GPS data from licensed third-party datasets and aggregation to 'core places'
median_dwell	Median minimum dwell time in minutes	Processing of raw GPS data from licensed third-party datasets and aggregation to 'core places'

B. Network Distances

To incorporate network distances into HDBSCAN (Section 3.4), a (network) distance matrix was computed for each iteration (county) of the algorithm. To extract each matrix, a local implementation of the Open Source Routing Machine (Luxen and Vetter, 2011) was set up and accessed using the ‘osrm’ R package (Giraud, Cura and Viry, 2020). The output was a matrix containing the network distances from each point in the dataset to every other point in the dataset, which could then be used in HDBSCAN (*dist*).

C. Typology Variables

A detailed summary of each of the twenty variables used to build the retail centre typology, can be seen below in Table 17. The four composition variables were the proportions of different types of retail (e.g., comparison) in each centre, and the proportion of independent retailers (pct_independent) was calculated as the proportion of those ‘core places’ in centres identified as major brands (see Table 14), or occurring more than three times in Chicago (location_name).

Category diversity (category_diversity) was calculated as the proportion of distinct SafeGraph top categories in each retail centre.

As above in section A, several variables were extracted from the SafeGraph ‘weekly patterns’ mobility data, using data from a ‘typical’ week, to be used in the retail centre classification. Median distance travelled and dwell were taken as the median values of all retail places in a retail centre, and total visits was the sum of visits to all places in each centre. The geodemographic variables were extracted using population data from the US Census Bureau (2018), and the US geodemographic classification (Spielman and Singleton, 2015). To extract geodemographic and population estimates for the catchments of centres, approximate drive-time isochrones were derived at differing intervals based on the size and suggested catchment areas of different US shopping centres (ICSC, 2017), ranging from 3-minute to 15-minute drive-time catchments. The geodemographic variables were calculated as proportions of catchment population occupied by each geodemographic category (e.g., groupa_prop).

Table 17. Description and source of each variable used in retail centre classification.

Domain	Variable	Description	Source
Composition	pct_comparison	% clustered units defined as 'comparison' retail	SafeGraph 'core places'
	pct_convenience	% clustered units defined as 'convenience' retail	SafeGraph 'core places'
	pct_service	% clustered units defined as 'service' retail	SafeGraph 'core places'
	pct_leisure	% clustered units defined as 'leisure' retail	SafeGraph 'core places'
Diversity	pct_independent	% clustered units defined as independent	SafeGraph 'core places'
	category_diversity	proportion of distinct safegraph top categories	SafeGraph 'core places'
Size/Function	n	total number of retail locations in cluster	SafeGraph 'core places'
	roeck_score	centre morphology, defined by roeck score	SafeGraph 'core places'

	median_distance	median distance travelled to each centre (metres)	SafeGraph 'weekly patterns'
	groupa_prop	proportion of catchment population "Hispanic & Kids"	Spielman and Singleton (2015), US Census Bureau (2018)
	groupb_prop	proportion of catchment population "Low Income & Diverse"	Spielman and Singleton (2015), US Census Bureau (2018)
	groupc_prop	proportion of catchment population "Middle-Income, Single Family Homes"	Spielman and Singleton (2015), US Census Bureau (2018)
	groupe_prop	proportion of catchment population "Low Income, Minority Mix"	Spielman and Singleton (2015), US Census Bureau (2018)
	groupf_prop	proportion of catchment population "Old, Wealthy White"	Spielman and Singleton (2015), US Census Bureau (2018)
	groupg_prop	proportion of catchment population "Residential Institutions, Young People"	Spielman and Singleton (2015), US Census Bureau (2018)
	grouph_prop	proportion of catchment population "Wealthy Urbanites"	Spielman and Singleton (2015), US Census Bureau (2018)
	groupi_prop	proportion of catchment population "African American Adversity"	Spielman and Singleton (2015), US Census Bureau (2018)
	groupj_prop	proportion of catchment population "Wealthy Nuclear Families"	Spielman and Singleton (2015), US Census Bureau (2018)
Economic Health	total_visits	total centre visits	SafeGraph 'weekly patterns'
	median_dwell	median dwell time for each centre (minutes)	SafeGraph 'weekly patterns'

II. Integrating the *Who, What* and *Where* of U.S. Retail Center Geographies

Here a detailed overview of the variables used to create the retail centre typology in Section 4.5 is provided; in particular the datasets used and creation of each of the variables used in the four classification domains is outlined. As in the paper, the choice of variables is done in such a way as to align closely with those used in Dolega et al. (2021), with modifications made to suit the SafeGraph data, and to account for specific niches in American retail.

Domain 1: Composition

In this section, the creation of the variables used in the 'Composition' domain of the retail classification is outlined. The SafeGraph '*core places*' dataset contains category aggregations for every place in the US, comprising a higher-level 'top_category' and lower level 'sub_category'. The purpose of these variables in the dataset is to ascribe a category aggregation for each place that describes the purpose of the place (e.g., jewellery store). In Dolega et al. (2021), the authors utilised the raw categories in the LDC dataset to compute variables that directly investigated the composition of retail centres. However, in the SafeGraph dataset both the 'top_category' and 'sub_category' variables contained high numbers of distinct values (193 and 388 respectively), thus creating variables for each distinct value of these would likely not be useful.

Thus, this approach involved creating an additional aggregation with only sixteen values that could be used to extract variables capturing the overall composition of the retail centres, based on the categories used in the original approach (Dolega et al., 2021). This aggregation was created by grouping together the places by 'sub_category', based on sensible overlaps in terms of the goods/services that they offered. Examples of how this aggregation was performed can be seen in Table 18 below.

Table 18. Aggregation of SafeGraph sub_category to higher-level aggregation for calculation of composition variables.

Aggregation	Example 'sub_category'
ClothingandFootwear	Women's clothing stores; jewelry stores; shoe stores
DIYandHousehold	Furniture stores; floor covering stores; paint and wallpaper stores

Electrical	Electronics stores; electrical apparatus and equipment; consumer electronics and appliances
Recreational	Musical instrument and supplies stores; art dealers; hobby, toy and game stores
Chemist	Pharmacies and drug stores; drugs and druggists' sundries; all other personal care stores
CTNandGasoline	Confectionery and nut sotres; gasoline stations and convenience stores; tobacco stores
FoodandDrink	Convenience stores; supermarkets and other grocery stores; fruit and vegetable markets
GeneralMerchandise	General merchandise stores
Bars	Drinking places (alcoholic beverages)
Restaurant	Full-service restaurants; cafeterias, grills and buffets; snack and nonalcoholic beverage bars
FastFood	Limited-service restaurants
Entertainment	Casino hotels; amusement and theme parks; historical sites
Fitness	Fitness and recreational sports centres; sports teams and clubs; golf courses and country clubs
ConsumerServices	Consumer lending; postal service; beauty salons
HouseholdServices	Automotive repair and maintenance; home and garden equipment repair and maintenance; limousine service
BusinessServices	Commercial screen printing; investment advice; direct mail advertising

Once the aggregation had been performed, it was possible to extract the individual variables for the composition domain. This, as seen in the equation below, was completed by calculating the total number of units in each aggregation category – e.g. the total number of clothing and footwear stores (nClothingandFootwear) in each retail centre, relative to the total number of units in the centre itself (nUnits). The resultant variables reflected the proportion of retail centre units occupied by each of the sixteen categories (propClothingandFootwear), as outlined for an example in the equation below, and listed in full for all variables in Table 4.

$$propClothingandFootwear = (nClothingandFootwear / nUnits) * 100$$

Domain 2: Diversity

In this section the creation of variables used in the 'Diversity' domain of the retail classification is outlined. The SafeGraph 'brands' column, which contains an associated brand name for many of the SafeGraph retail places, was used to identify 'chain' and 'independent' retailers directly, with those without an associated 'brand' were labelled as independents. The chains, those with an associated 'brand', were then split further, with those retail places associated with a 'brand' with less than 50 stores across the US being labelled as 'small multiples' and those with 50 or more being 'national chains'.

These definitions of 'independent', 'small multiple' and 'national chain' were then used to derive three variables reflecting the occupancy of the retail centre. For each variable, the total number of each was calculated relative to the total number of units in each retail centre, before being converted to a proportion. An example of this for national chains can be seen in the equation below.

propNationalChain = (nNationalChain/ nUnits) * 100

To investigate the prevalence of the most common and thus 'popular' brands, three variables looking at the most 'popular' brands in three different types of retailing were calculated; comparison, convenience and leisure. First, the aggregations from Table 18 were aggregated further into 'comparison', 'convenience' and 'leisure', as seen below in Table 19, based on sensible overlaps between the two. The justification for doing so was that Dolega et al. (2021) used similar measures to explore diversity levels in contrasting functions of retail (e.g., convenience vs leisure).

Table 19. Identification of comparison, convenience and leisure aggregations from Table 16.

Domain 2 Aggregation	**Domain 1 Aggregation**
Comparison	ClothingandFootwear; DIYandHousehold; Electrical; Recreational
Convenience	Chemist; CTNandGasoline; FoodandDrink; GeneralMerchandise
Leisure	Bars; Restaurant; FastFood; Entertainment; Fitness

Once these aggregations had been performed, the most popular 'brands' in each were identified by extracting the top 100 most frequently occurring 'brands' in 'comparison', 'convenience' and 'leisure' retail. Once these lists had been compiled, the proportion of total retail centre units occupied by popular (comparison, convenience, and leisure) 'brands' was calculated, as seen for an example in the equation below, and listed in full in Table 4 of the manuscript.

propPopularComparisonBrands = (n.PopularComparisonBrands / nUnits) * 100

A diversity index for retail (comparison, convenience, leisure) and services was also constructed both at the national and local level, as in Dolega et al. (2021). These four variables (nationalRetailDiversity, localRetailDiversity, nationalServiceDiversity, localServiceDiversity), were assembled by exploring the diversity of SafeGraph 'top categories' in retail centres, relative to the total number of available 'top categories' nationally and at state-level (for the local indices). Examples for retail diversity at the national and local levels can be seen in the equations below.

$$\text{nationalRetailDiversity} = (\text{n.RetailCategories} / \text{n.NationalRetailCategories}) * 100$$

$$\text{localRetailDiversity} = (\text{n.RetailCategories} / \text{n.LocalRetailCategories}) * 100$$

Domain 3: Size and Function

In this section, the creation of variables used in the 'size and function' domain of the retail classification is outlined. The total number of units in each retail centre (nUnits) was derived by calculating the total number of SafeGraph retail places in each retail centre, and the total number of buildings (nBuildings) was calculated in the same way, instead using the building geometries associated with the retail places. The area of the retail centre (area) was calculated using the st_area() function from the sf R package, providing a measure of area for the retail centres in km^2. The retailDensity variable was obtained by dividing the total number of retail buildings in the centre, by the total area of the centre itself, as in the equation below.

$$\text{retailDensity} = \text{nBuildings} / \text{area}$$

The morphology/linearity of the retail centres (roeckScore) was calculated by computing the roeck degree of compactness for each retail centre, where the area of the retail centre (area), is divided by the area of its smallest enclosing circle (areaEnclosingCircle), as in the equation below. Higher values of roeckScore indicate a more linear, less compact retail centre.

$$\text{roeckScore} = \text{area} / \text{areaEnclosingCircle}$$

Median distance travelled (medianDistance) was calculated using the SafeGraph '*weekly patterns'* data. The *'weekly patterns'* data provides a variable for all retail places that indicates the median distance travelled in kilometres (median_distance), to that retail place by consumers. To calculate the variable (medianDistance), the median value of median_distance for all retail places in a retail centre was taken, using data from July 2021.

The residentialDensity, retailEmploymentDensity and roadDensity variables were extracted using the Smart Location Database, from the US Environment Protection Agency (EPA, 2018). The data comes at the census block group level, thus those census block groups within the boundary of each retail centre were extracted, to gather information about the urban fabric of the retail centre itself. The individual Smart Location Database indicators of interest were extracted, and each is measured slightly differently – D1a (household units per acre), D1c8_Pub(retail jobs per acre) and D3a (road network density, considering multi-modal transport, pedestrian links and intersections). To calculate these at the retail centre level, all census block groups within the retail centre boundary were extracted, and the median value for the three SLD indicators were taken, as listed in Table 4.

The existence of 'anchor' stores was measured similarly to in the original study (Dolega et al., 2021), by identifying the highest grossing comparison retailers, using official SafeGraph documentation (SafeGraph Inc., 2022), and the highest grossing general merchandise retail brands, using information from the National Retail Federation (2019). The list of each can be seen below.

- Comparison retailers: JCPenney, Bloomingdale's, Macy's, Nordstrom, Saks Fifth Avenue and Neiman Marcus.
- General merchandise/highest grossing: Walmart, The Kroger Co, Costco, Walgreens, Target.

To calculate the propAnchor variable, the total number of Anchors, from the list above, in each retail centre (nAnchor) were calculated, before converting to a proportion relative to the total number of units in the centre (nUnits), as seen in the equation below.

$$propAnchor = (n.Anchor / nUnits) * 100$$

To extract propDiscount, those 'brands' present in the SafeGraph retail places dataset and on the official list of top 50 'Discount stores of the US' were extracted (Wikipedia, 2022), and the proportion of these in each retail centre was calculated, as in the equation below. To extract the premium brands (propPremiumBrand), those 'brands' present in the SafeGraph retail places dataset and on the official ranking of top 50 'Luxury and Premium Retail Brands' (Brand

Finance, 2022) were identified, and calculated the proportion of these in each retail centre, as in the equation below.

$$propDiscount = (nDiscount / nUnits) * 100$$

$$propPremiumBrand = (nPremium / nUnits) * 100$$

Domain 4: Economic Health

In this section an overview of the variables used in the 'economic health' domain of the retail classification is provided. The total number of visits (totalVisits) was calculated by summing the total weekly visits to each retail place within each retail centre, again using the SafeGraph '*weekly patterns'* data from July 2021.

To identify catchment characteristics (totalPopulation, medianIncome, medianUnemployed), approximate fixed-ring buffers were derived, varying in size according to the thresholds set out in the ICSC classification for different sized [shopping] centres (ICSC, 2017). Drive-time catchments were not utilised due to unnecessary additional computational cost. These catchments were used to extract the median income and median proportion of unemployed people at census block group level, using the 'tidycensus' and 'tigris' R packages, through calculation of median values from those census block groups falling within each catchment.

The total number of competing destinations (nCompeting), was derived by extracting the number of those similar-sized centres within the catchment, as set out by ICSC thresholds, and using the fixed-ring buffers used to extract catchment characteristics. The balance of retail to service (retailService) was calculated by computing the proportion of retail units (integrating comparison, convenience, and leisure places) and the proportion of service units, both relative to the total number of units in a retail centre (nUnits), before subtracting the latter from the former.