Spatial Pattern Mining of Tech Clusters of Dynamics and Industry Mix Based On Quantitative Methods in England Area, UK

Zeqiang Fang

CASA0012, MSc Spatial Data Science and Visualisation Dissertation

Supervisor: Dr Max Nathan

Repository: https://fang-zeqiang.github.io/CASA0012-Dissertation/

This dissertation is submitted in part requirement for the
MSc in the Centre for Advanced Spatial Analysis,
Bartlett Faculty of the Built Environment, UCL

Word count: 10,050

2021-08-23

Abstract

Since tech clusters play an increasingly important role in economic contribution around the world, the analysis of what factors can affect firm dynamics has become a relatively new topic in economic geography research. England's tech industry is a vital aspect of the global economy, and the UK government has included it in numerous investment programs. Therefore, studying the distribution of science and tech clusters, regression analysis and spatio-temporal pattern mining methods to comprehensively understand the relationship among industrial concentration, firm density and dynamics will help to provide appropriate local policy recommendations. Despite limitations of the research framework, regression models and spatial autocorrelation methods have good reproducibility. Based on quantitative methods, it is found that the higher the degree of industrial diversification and the greater the density of local tech companies, the higher the entry rate of tech firms. Simultaneously, the visualisation of spatio-temporal patterns also indicates the transfer of hot spots of firm entry. These results are conducive to optimising the government's investment allocation in England tech industry.

Declaration

I, Zeqiang Fang, hereby declare that this dissertation is all my own original work and that all sources have been acknowledged. It is 10050 words in length

Acknowledgements

This paper would not have been possible without the help of others. First of all, I am sincerely grateful to my supervisor, Prof. Max Nathan, for his patient guidance and enthusiasm support throughout my dissertation project. Prof. Max Nathan not only provide me with much inspired feedback, but also offer me much care of the difficulties I encountered in thesis writing, both personally and academically.

Contents

List of Figures

List of Tables

Abbreviations

Term	Abbreviation
Travel to Work Area	TTWA
Herfindahl-Hirschman Index	HHI
Local Indicators of Spatial Association	LISA
Ordinary Least Squares	OLS

Chapter 1

Introduction

1.1 Background and Motivation

Silicon Valley and other tech clusters are critical for modern innovation, company competitiveness, and economic performance in this digitalising world(Kerr, 2020). In terms of the development of high-tech industries, UK is becoming increasingly attractive to international investors. And the scale of VC investment in the UK had reached the third level in the world in 2020, even if it is severely affected by the covid-19 epidemic (Tech Nation, 2021). The UK tech startup and scaleup ecosystem are also valued well. Investment and relative infrastructures are friendly for tech companies (ibid). In fact, high technology companies are often investigated from the dimension of travel to work areas in many fields of research. For example, some reports found that, between 2007 and 2014, the number of creative tech enterprises grew faster than the overall company population in more than nine out of ten of the UK's 228 travel to work area geographies (Mateos-Garcia 2016).

There are clustering effects existing among the England tech firms. Some research teams such as Baptista (1999) conducted growth model research and dynamic model mining on regional clusters of technology industries and found that the growth rate of the sub-category business of a specific industry in the cluster may be higher than the industry average and a certain agglomeration phenomenon. Some teams also had done research on employment changes in the corresponding technology research and development industry. Also, the employment growth of high-tech small and medium-sized enterprises (SMEs) seems to depend on the initial level of clusters, and the impact of this clustering effect on employment growth varies greatly in geographic scope (Fingleton et al., 2004). Although there are obvious differences in clustering characteristics between the technology industries in the UK and other areas, the cluster effect of the technology industry in the UK did not seem to be weaker than that in the United States (Baptista and Swann, 1999).

The dynamics characters of clusters can be measured by their entry pattern. Many patterns of industrial dynamics appear to be affected by the process of knowledge creation, gathering, and destruction. This promotes the admission of new businesses, the long-term coexistence of incumbents and newcomers, and their selected or combined withdrawal (Krafft, 2004). In terms of industry clustering pattern and policy formulation, Nathan and Rosso (2015), in response to the development needs of governments for information and communications technology sectors, utilise big data resources to conduct an innovative alternative analysis for all active information and communications technology manufacturers in the UK. This can help researchers to better understand the current condition of the UK's information and communications technology industry. Meanwhile, the approach of spatial autocorrelation analysis is useful for revealing the structure and patterns of economic geographical variables. It can be used to identify not only the country's overall spatial patterns but also specific micro-locations(Stankov and Dragićević, 2015). Our research can employ these methods and background information to explore the technology industry dynamics and obtain sufficient argument evidence.

1.2 Research Question and Objectives

In order to contribute to this field research, the main purpose of this research is to explore the spatial pattern of tech clusters characteristics, introduce quantitative methods to explore the relationship between cluster dynamics, industrial concentration and density. The research will focus on the following main questions:

> To what extent will industry mix and firms density affect tech clusters' dynamics in England area
>
> How does tech clusters' dynamics spatio-temporal pattern change in England area from 1998 to 2018?

The overall experiment will be carried out around the above-mentioned central question. Breaking down the above central issues, the objectives to be achieved at each stage of this research are divided into 5 points as follows

1. Review the relevant literature on the use of regression models to identify the relationship between company clusters and find the empirical research which employs Moran's index(Moran's I), local indicators of spatial association(LISA) and hot-spot analysis based on the Getis-Ord Gi* method
2. Processing and filtering data which can represent the dynamics and related characters of each tech cluster area in England from 1998 to 2018
3. Exploring the relationship between tech cluster dynamics, industry concentration and company density through trend analysis and regression models
4. Based on the mining and visualisation of related spatial indicators, understand the spatial autocorrelation of relevant indicators and the trend in the temporal dimension among various travel to work areas (TTWAs) within the England area.

5. Combine with actual policies to provide guidance and recommendations at the company level and the government level based on regression results and spatial model mining results

1.3 Report Structure

This research paper contains 6 chapters. Chapter 2 will first compare and summarise the current research status of tech cluster dynamics in the UK and other parts of the world. In Chapter 3 methodology, the research scale and object will be explained. Data processing process and index data summarised in different dimensions will be presented. Research design combining multiple quantitative methods will also be included in this chapter. Chapter 4 will analyse the changes in distribution and explore the influence relationship among variables from the overall to the partial. In addition, this chapter will compare the global and local indices for spatio-temporal analysis. Chapter 5 will discuss the experimental results based on the actual relevant policies of the England authority, and give investment suggestions at the government and enterprise levels for the development of the tech industry and discuss the limitations of the research framework. Finally, Chapter 6 will summarise key findings to respond to the research questions and make recommendations for future study and planning strategies.

Chapter 2

Literature Review

2.1 Industrial Cluster and Tech Cluster

A cluster is described as a geographic concentration of businesses that operate in the same industry. High rates of entry and exit are often correlated with growth in employment and productivity, so these questions are of interest not only to academics but also to policymakers (Fritsch, 2011). Industry clusters form as a result of the production benefits of local specialization mixed with cross-regional trade. Three forces of what is now known as agglomeration economies were discussed previously: knowledge spillovers, labour market pooling, and customer-supplier interactions. While most Marshallian force studies have focused on industrial contexts, they also apply to tech clusters, albeit in slightly different ways (Ellison et al. 2010). The definition of tech clusters is mentioned in Kerr's research(2020) as given below.

> we define "tech" clusters to be locations where new products (be they goods or services) and production processes are created that impact multiple parts of the economy.

In identifying Information and communications technology (ICT) relative economic activities, They find ICT employment shares nearly double the usual estimates using their unique approach even if this conclusion is more speculative(Nathan and Rosso, 2015). The definition and tech relative industry information can give a brief understanding of our research objects.

2.2 Industry and Cluster Dynamics

Industrial dynamics is still a relatively new subject of research. Its main focus is on the evolution of industries, with an emphasis on company entry, growth, and exit(Malerba, 2006). Industrial dynamics, in contrast to the neoclassical approach to industries (also known as industrial organization), focuses primarily on phenomena that are intrinsically dynamic and deals with them in an interdisciplinary manner such as economic geography(Carlsson 1987;Fritsch, 2011).

In industrial clutser analysis, concepts of anchor firms for clusters are developed by some researchers. This thesis indicates that big corporate headquarters can help a specific area attract many other firms to form a new cluster and this anchor firm is usually recognised as the basis of the cluster (Feldman, 2003; Porter, 1998). For the cluster dynamics, the cluster has a strong influence on the company's entry behaviour. The relevant research team found that the entry rate increases as the cluster size increases (Frenken et al., 2014). Most of the research focuses on the correlation between industry dynamics and localisation but there is little extensive quantitative research on the relationship between industry concentration and entry pattern(Furman et al., 2002; Shaver and Flyer, 2000).

Relative researchers found that industry clustering not only can increase firm entry but also firm exit rates. This implied that clusters could emerge and exist because they provide entry opportunities but they do not necessarily generate Marshallian economies that increase firm survival (Boschma, 2015;Alcacer and Chung, 2007). Some teams investigate cluster emergence and evolution over the long term by looking at how clusters influence entry, exit, and growth via localisation economics(Kogut,1991). Besides, Clementi and Palazzo(2016)found there is a cluster effect on the whole of England area. The arrival of new businesses has a significant impact on the overall dynamics. The cluster's new entrants will grow in size over time, resulting in broader expansion and overall influence. This research framework can be referred to our quantitative measurement of cluster dynamics.

2.3 Industry Concentration

In competition law, antitrust, and technology management, the Herfindahl–Hirschman Index (HHI) is a widely used economic concept. It's a metric for a company's size in proportion to its industry, as well as a barometer for how competitive it is (Liston-Heyes and Pilkington, 2004). On average, businesses in big cities are more productive. The natural advantages of localisation may enhance corporate choice and agglomeration economies. Large cities encourage competition by enabling only the most productive individuals to stay, fostering contact and increasing productivity(Porter and Stern, 1999).

Tremblay et al.(2005) have studied the relationship between the brewing industry and the entry of enterprises from a macro and micro perspective. They found that regional technological changes will lead to an increase in industry concentration and macro-fields. Excessive industrial prosperity and high concentration will lead to a decline in business entry and a decline in survival rates(ibid). The Herfindahl–Hirschman index is also used to highlight the dynamic patterns of the sustainability sector, which is crucial for understanding how startups have applied the concept of sustainability(Kwon, 2020).

2.4 Other Factors and Dynamics Pattern

The density of regional firms may have a two-sided impact on the entry pattern. Because each firm has the ability to bring in new entrants, firm density has a positive effect on entry rates in the early stages of an industry. Due to severe market rivalry, higher firm density levels, on the other hand, become a barrier to entry as the industry expands and flourishes (Boschma, 2015). Meanwhile, empirical research shows that the relationship between business size and the likelihood of survival is influenced by technology and the stage of the industry life cycle (Agarwal and Audretsch, 2001).

Studies have also shown that entry may be less about radical innovation and more about filling strategic niches in established businesses that are nonetheless technologically intensive, neutralizing the impact of entry size on the likelihood of survival(ibid). There is also evidence that a negative relationship between firms' age and growth over time is because more and smaller firms such as factories create more jobs than some large companies(Liu et al., 1999). For firms' growth and failure, Moreno's research(2000) found that the failure rate and growth rate of companies decline with the growth of scale and age. Regression-to-the-mean is not an important factor in the negative correlation between the size and growth of surviving companies. And he also offers empirical evidence of firm failure rates as well as the mean of the distribution of realized growth rates.

2.5 Conclusion

Overall, by reviewing the past literature on industrial clusters and enterprise dynamic models, it can be found that industrial dynamics is still a relatively novel research topic in the UK. Usually, the characteristics of research clusters can be conducted in an interdisciplinary way such as economic geography. As a region with obvious cluster effects, the UK's industrial clusters can increase the entry of enterprises. And there is also a reference for the relevant literature on the quantitative measurement method of the entry index. In terms of industry concentration, a certain degree of industrial concentration in the short term may attract companies to enter, that is, industry prosperity within a certain period of time. But excessive entry will also cause industry shocks. Through the research review of the dynamic pattern of enterprises, it is also found that the regional density level, the age of enterprises or clusters and etc. have an influence on the entry pattern of enterprises.

Chapter 3

Methodology

3.1 Research Framework

The study area is 149 travel to work areas in England area (out of 228 in the UK). The areas in blue colour is the main research area (Figure 3.1). In this study, a data set containing all companies in UK will be cleaned and attribute selected, and technology companies will be identified and screened according to the classification and definition of technology companies on the official website of the British government. Before the quantitative study, this study is based on time and The spatial dimension counts the number of technology companies, and calculates the dynamic indicators of enterprise clusters and industrial combination indicators in a specific year and a specific region. Then this study conducts multiple regressions, univariate variables Moran index testing to conduct spatial quantitative research, and finally combines Qualitative spatial pattern trend research on the spatial changes of indicators in three different time periods(1998, 2008 and 2018).

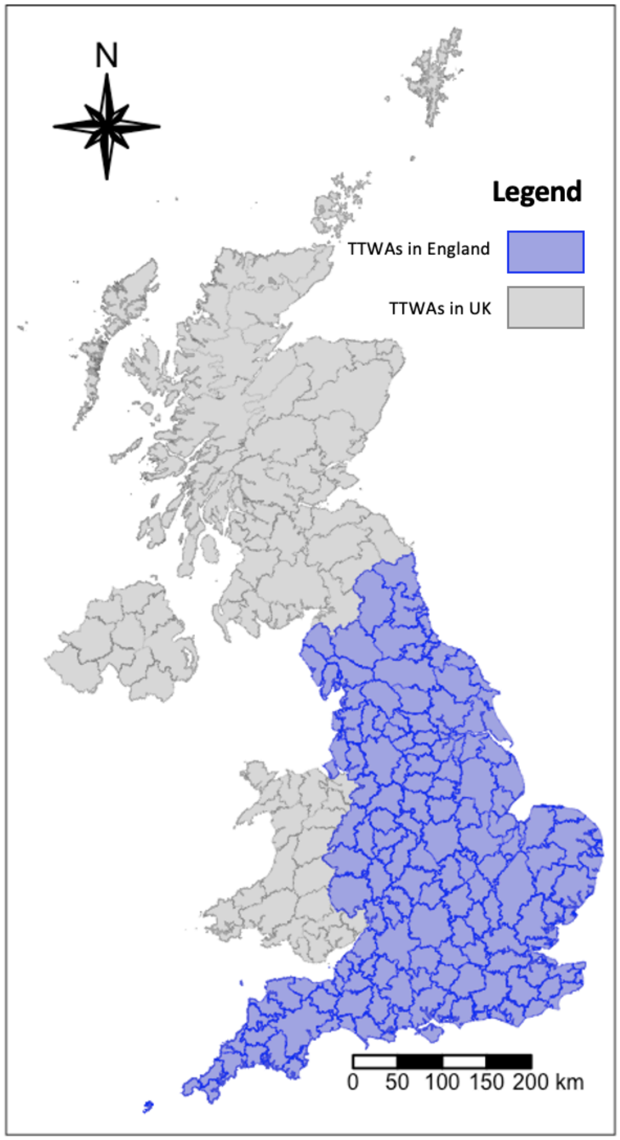

Figure 3.1: The Map of Travel to Work Areas (TTWAs) within UK

3.2 Data Source and Processing

This raw dataset is collected from the core company data from Open Corporates master company database (Open Corporates, 2018). And the size of dataset accounts for 15 GB which is handled with `read_stata` and `get_chunk` function to read large data file in chunks, then increasing the reading speed. The "primary uk sic 2007" identification field is the basis of industry finding and the "birth year" is the key to measure dynamics variables . All rows whose these two values are empty are removed 17% incorperate date is missing and sic code is complete). The details of data processing can be seen in the reproducible analysis reports in Appendix B.

3.3 Identifying Tech Clusters

For the identification of science and technology companies, this study introduces the main 2007 sic code table to judge the science and technology industry, referring to the classification method of the Science and Technology Classification data set on the ons.gov.uk website. In order to better identify science and technology companies for the UK. The economic contribution is officially based on the 2007 British Standard Economic Activity Classification, different data sources were combined to classify and label science and technology companies (Office for National Statistics, 2015).The 5-digital sic code can identify the economic activity classification of the technology industry in the finest granularity. The statistics bureau provided a comparison table of sic codes for the classification of science and technology industries, which can help this research to better identify technology companies from the big data of UK companies (3.1).

This research refers to the science and technology classification table provided by the government. The technology indicator is used to position the technology industry of all industries, and a total of 168 sic codes for the technology industry in 2007 were obtained, accounting for about 16% of all industry categories in the UK, including 5 industry categories such as Digital Technologies, Life Sciences and Healthcare,

Table 3.1: Science and Technology Classification Sample (Part)

5-digit SIC07 code	**Science and Technology category**	**Science and Technology topic**
26110	Digital Technologies	Computer & Electronic manufacturing
58210	Digital Technologies	Digital & Computer Services
62090	Digital Technologies	Digital & Computer Services
63110	Digital Technologies	Digital & Computer Services
27510	Other scientific/technological manufacture	Electrical Machinery manufacturing
...	...	...
86230	Life Sciences & Healthcare	Healthcare services

Publishing and Broadcasting , Other scientific/technological manufacture and Other scientific/technological services, details of the classification form will be attached in the attachment. There are almost 20% firms in the raw data belonging to tech firms according to the method of category as mentioned above.

3.4 Data Selection

3.4.1 Time Range (Temporal Dimension)

The reason why this research choose tech firms which incorporate from 1998 to 2018 is because the number of technology companies established in the UK in the past 20 years is significantly higher than before 1998, as shown in figure 3.2.

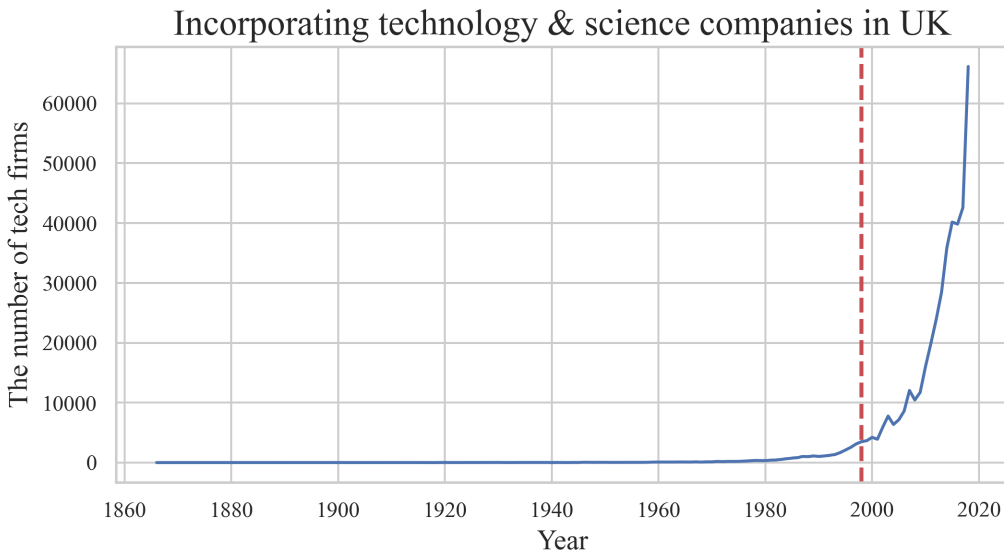

Figure 3.2: Incorporating technology and science companies in UK

3.4.2 Cluster Boundary (Spatial Dimension)

Travel to Work Locations is official statistics that capture local labour markets, i.e., areas where the majority (approximately 70%) of the people who live there work. These measures are based on responses to the 2011 and are used to define the TTWAs algorithmically. There are currently 228 TTWAs in the UK. When it is recognised that the activity of interest may be scattered across administrative boundaries such as local authority districts or NUTS areas, TTWAs are widely utilised in industrial clustering investigations (Prothero, 2021).

Relevant studies have shown that commuting across towns has become more common in England and Wales. People are not limited to living and working in the same administrative area. In addition, studies have found that low-skilled workers tend to rely on the public to work locally. Strong skill-oriented jobs are more dependent on cars, and they are also the main force for cross-city commuting. The technology industry has a greater demand for those strong skill-oriented workers which suggest that travel to work area(TTWA) is used as a cluster of technology companies compared to traditional administrative Region would be a good choice (Titheridge and Hall,2006). The University of Cambridge had funded some researchers to undertake the Wisbech 2020 Vision project to analyse the current problem, mining the potential future space for employment growth with alternative macro-economic scenario to help drive a high value-added growth plan in the local area (Burgess et al., 2014)

This geographical division can better reflect the relationship between population, company and work. In terms of geographic research, related researchers have combined the Business Structure Database (BSD) from the Office for National Statistics (ONS) and industry classification methods to use the ONS geographic area of TTWA to survey the commuting patterns and labour market of the population in 2011. Most scholars state that this might be an effective measure for the research of industrial clusters at the sub-regional level (Mateos-Garcia 2016).

3.5 Quatitative Analysis and Methods

3.5.1 Dynamics Measuring Index

Firm entry is the result of the interaction between the characteristics of an actor, on the one hand, and the surrounding environment, on the other hand (Frenken, Cefis and Stam, 2014).

To measure the degree of dynamic change of a cluster, it is necessary to calculate the entry rate of the technology cluster. Brandt used the enterprise's entry rate and exit rate to measure the dynamics attributes of a company (2005). This research refers to the researcher's calculation method. The number of enterprises entering and quitting a sector as a percentage of the total number of firms in the same sector in a given year is used to compute entry rate in this tech cluster dynamics research(ibid). The calculation method is as follows.

$$Entry\ Rate_{i,t} = \frac{Incorporating\ Firms_{i,t}}{Total\ Firms_i}$$

where i means location(travel to work area), t means year

3.5.2 Industry Mix Measuring Index

It is necessary to calculate the Herfindahl-Hirschman Index or location quotient of the technology cluster to measure the degree of industrial concentration in a cluster area. However, only the former method is used because the employment data corresponding to the corresponding region year is missing; Here is a reference to the quantitative method of industrial concentration of Chao Lu's research team (2017). HHI is calculated by squaring the market share of each competing company and then adding the results, where the market share is given in the form of scores or points (ibid). Increases in the Herfindahl index generally indicate a decrease in competition and an increase of market power, whereas decreases indicate the opposite (Hall et al., 2009), as the calculation method shown below.

$$HHI_{\ i,t} = \sum_{j=1}^{N} \left(\frac{Tech\ Firms_{j,i,t}}{Total\ Tech\ Firms_{i,t}}\right)^2$$

where

$HHI_{i,t}$ means Herfindahl-Hirschman Index in location i and year t

N is the overall number of individual tech firm contained.

k represents the kth industry in location i and year t.

$Tech\ Firms_j$ is the number of j-th individual tech firm

$Total\ Tech\ Firms$ represents the number of total tech firms in a specific location and year.

3.5.3 Regression

In order to quantify the impact of industry concentration and company density in the cluster on the cluster dynamic variable (entry rate), the Herfindahl-Hirschman

Table 3.2: Independent Variables Selection for Regression

Independent Variables	Type	Description
Density	numeric	The density of tech firms
Herfindahl-Hirschman Index	numeric	The index measuring the industry mix
Year	category	Years from 1998 to 2018
Location	category	Top ten clusters of companies in England

Index and company density corresponding to the year and location(i.e. ttwa) in the data are used as independent variables (Table 3.2). And the variable of the numeric data type of entry rate is regarded as the dependent variable

Dynamics variable(entry rate) is measured empirically by the ratio of the number of new firms at the end of a period to the total number of firms at the beginning of the period(Hause and Du Rietz, 1984), unless otherwise stated. The equation regression model is given:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \sum_{i=3}^{4} b_i x_i + \epsilon_{i,j}$$

where

y means entry rate in the year and location x_1 means the year, a dummy variable x_2 means location, i.e. travel to work area(TTWA) x_3 means the density of tech firms in a specific year and location x_4 means Herfindahl-Hirschman Index in in a specific year and location $\epsilon_{i,j}$ is a random error.

3.5.4 Global Autocorrelation (Moran's I)

One of the most popular and often used measures of spatial autocorrelation is Moran's Index. Based on the locations and values of the feature, the Global Moran's I tool analyzes the pattern of a data set spatially and decides if it is scattered, clustered, or random. The program calculates the Moran's I Index value, as well as the z-score and p-value, to statistically validate the Index. It is computed using the formula below (Kumari and et al., 2019).

$$Moran's\ I = \frac{n}{S_0} \frac{\sum_{x=1}^{n} \sum_{y=1}^{n} w_{x,y} z_x z_y}{\sum_{x=1}^{n} z_x^2}$$

$$S_0 = \sum_{x=1}^{n} \sum_{y=1}^{n} w_{x,y}$$

where z_x stands for deviation of an attribute from its mean $(x_i – X)$ for feature X, $w_{x,y}$ is the spatial weight among feature X and Y, n is the total number of features and So is the summation of all spatial weights

The statistic's z_j score is given below

$$z_x = \frac{I - E[I]}{\sqrt{V[I]}}$$

where

$$E[I] = \frac{-1}{(n-1)}$$

$$V[I] = E[I^2] - E[I]^2$$

The Moran's Index(I) has a range of values from -1 to 1. Index value 1 indicates that the observed pattern is spatially clustered. On the other hand value -1 indicates scattering or dispersion. Moran's I assign a value of near or equal to zero to the lack of auto correlation. The z-score and p-value of the Index are only used to form final judgments regarding the observed pattern. Besides, local indicators of spatial association will be applied to measure the high-high and low-low clusters(Anselin, 2010). According to research method of Kumari's team(2019), the Moran's I of three time period will be calculated to analysis the dynamics change in spatial and temporal aspects.

3.5.5 Hot-Spot Analysis (Getis-Ord Gi*)

Only positive spatial autocorrelation is taken into account by the local Getis statistics, which allows for discrimination between clusters of similar values that are high or low relative to the mean. The Getis–Ord local statistics are calculated as follows by Getis and Ord (2010):

$$G_j^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{[n \sum_{j=1}^{n} w_{i,j}^2 - (\sum_{j=1}^{n} w_{i,j})^2)]}{n-1}}}$$

where x_j is the attribute value for feature j, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features and:

$$\bar{x} = \frac{\sum_{j=1}^{n} x_j}{n}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - (\bar{x})^2}$$

The Gi* statistic returned for each feature in the dataset is a z-score (Getis and Ord, 2010). For statistically significant positive z-scores, the larger the z-score is, the more intense the clustering of high values (hot spot). For statistically significant negative z-scores, the smaller the z-score is, the more intense the clustering of low values (cold spot).

3.6 Limitations

In term of missing value in the original dataset, there are almost 99.8% missing value in dissolution date. This means that most companies do not have a dissolution date, which might not mean that some companies survive, nor can it accurately reflect the company's exit numbers in a specific year and region. Moreover, the company's incorporation date data (including firms birth year) is also missing about

17.26%. These two difficulties make the data after cleaning process may have the risk of insufficient accuracy in representing the dynamics of the industry. The model prediction after fitted to this dataset might not reach the same situation as the real world level.

England travel to work areas might not be a data-driven method to identify the clusters. This spatial boundary data is not an up-to-date dataset because it is lastly updated in 2015. Changes in commuting have pushed TTWA boundaries further since their inception in the 1980s. With the passage of time, the boundaries of TTWA will be further changed in the future due to changes in commuting methods(Ozkul,2014) which might make this research slightly inappropriate in the future application.

The value of this statistical data(Herfindahl-Hirschman Index) for identifying monopoly development, on the other hand, is directly dependent on the precise definition of a market (Kwoka, 1977). For instances, geographical considerations might influence market share. This dilemma can arise when there are nearly equal market share of tech businesses in a given sector, but they each operate exclusively in distinct regions of the travel to work area, resulting in each firm having a monopoly inside the specific marketplace in which it conducts business, which might make it more difficult to measure the industry mix in a specific location and year. Furthermore, one IT firm may control as much as 70% of the market for a certain area of the digital industry (i.e. the sale of one specific equipment). As a result, that company would have a near-total monopoly on the manufacturing and sale of that commodity.

3.7 Ethical Statement

The data for this project comes from OpenCorporates, a firm that aggregates company-level data from around the world(https://opencorporates.com/). In this case, OpenCorporates have taken data from the UK Companies House register (https://www.gov.uk/government/organisations/companies-house). As detailed by

Nathan and Rosso (2015), all limited companies in the UK need to registers with Companies House when they are set up, and provide annual returns and financial statements. These include details of directors and company secretary, registered office address, shares and shareholders, as well as company type and principal business activity. Thus, all the data used here is already in the public domain.

The research objectives are tech firms in the UK for this project and the individual data will not be collected and measured in this project. For issues of deanonymisation or privacy, traceable information such as the real companies name and ID will not be utilised in the research. The raw data will be cleaned and filtered by several key variables include industries instead of the company's name or other sensitive information before doing the research. Through data cleaning, pre-processing, desensitisation or other processing methods, the risks of damage to company interests (such as social reputation, economic benefits and etc.) will be mitigated to an as low as possible level in the research process. Besides, this project will not cause discrimination of industries or job categories. The final analysis results, such as the different industry concentration in each region, will not deepen some people's stereotypes and prejudices about the region (This content will be fully discussed in the project discussion section).

It is necessary to point out and declare the objectivity of the analysis and the non-absoluteness of the results in the disclaimer. Consider the feelings of people and governments in different parts of the UK, this research will prevent the influence of personal preferences and subjective emotions. Last but not least, the leakage of companies' name and information will be protected. For example, in the reflection of the results section of academic research, the name and related information of the companies that moved may be revealed. Although this information may be open to the public you need to know that this information may be used by people with other ulterior motives. This project will desensitise the company name and information at the stage of chart presentation, such as using A, B, and C to replace them to achieve this purpose.

Chapter 4

Results

4.1 Spatial and Temporal Distribution

In terms of three variables' change of overall England level,it can be found that the average firms' entry rate and density in England increase significantly(Figure 4.1). There was a growth in firms entry rate from 0.01(1998) to 0.14(2018) and the statistics of density climb from 0.02 to 0.43 during these decades (Table 4.1). However, the Herfindahl-Hirschman index decreased by less than 0.1 in 2018 which is 1/3 times than that in 1998. Overall, it can be seen that the dynamics increased but the industry concentration reduced.

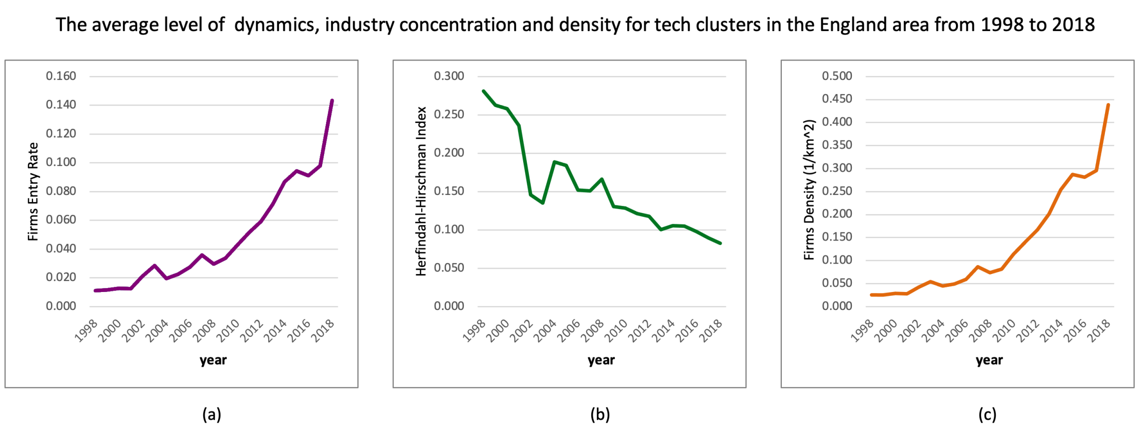

Figure 4.1: The average level change of three variables in two decades

Table 4.1: The average level change of three variables in two decades

Year	**Entry Rate**	**Herfindahl-Hirschman Index**	**Density (1/km^2)**
1998	0.0111338	0.2811410	0.0251467
1999	0.0116754	0.2625985	0.0255016
2000	0.0127544	0.2580325	0.0288208
2001	0.0124221	0.2361455	0.0279131
2002	0.0212760	0.1460459	0.0420549
2003	0.0285419	0.1350075	0.0540107
2004	0.0195123	0.1886668	0.0449568
2005	0.0225973	0.1841296	0.0492084
2006	0.0274033	0.1521901	0.0592952
2007	0.0358175	0.1507693	0.0862518
2008	0.0295100	0.1662619	0.0737782
2009	0.0336153	0.1305273	0.0817783
2010	0.0425260	0.1285513	0.1134154
2011	0.0516425	0.1212957	0.1404650
2012	0.0593572	0.1179875	0.1664194
2013	0.0712224	0.1006216	0.2015662
2014	0.0867586	0.1056914	0.2538923
2015	0.0943852	0.1048371	0.2869250
2016	0.0912387	0.0980931	0.2812983
2017	0.0979282	0.0894585	0.2954844
2018	0.1432780	0.0827082	0.4387067

4.1.1 Tech Cluster Dynamics

From 1998 to 2018, the entry rate of each travel to work area (TTWA) in England will change drastically every decades. For example, in 1998 the darker colored areas were mostly around the Greater London area and the entire southern part of England(Figure 4.2). The overall entry rate in 2008 was higher than that in 1998 and showed a trend of new companies entering the north of England. In this year, some travel to work areas on the edge of England have more prominent entry rates, such as Malton, Minehead and Bridport. However, the overall entry rate in 2018 was about twice as high as the average in 2008. Moreover, ttwa with high entry rate is mostly concentrated in the north and the surrounding areas of Greater London. The overall trend can be witnessed as given by table 4.1 that there is an increasing trend in the overall level of entry rate from 1998 (average about 1.06%) to 2018 (average about 14.33%)

Finally, from the perspective of changes in temporal and spatial distribution, the increase in the entry rate of the travel to work area in the northern part of England is much larger than that in the southern part. The areas with high entry rates before 2008 were mostly scattered in other areas except for the Greater London area. However, close to after 2018, London and its surrounding areas, the central Manchester-Birmingham area and the southwest, such as Exeter, have a rising entry rate. In contrast, in 2008, the urban agglomerations north of Leeds had higher firm entry rates in tech industry(Figure 4.2).

4.1.2 Tech Industry Concentration

In terms of herfindahl-Hirschman index of tech industry in travel to work area of England. As given figure below, it can be seen that the Herfindahl-Hirschman index has a downward trend year by year (Figure 4.3). In the methodology module, this index is mentioned as a measure of industrial concentration. A decrease in the value represents an increase in industrial competition and a decrease in market

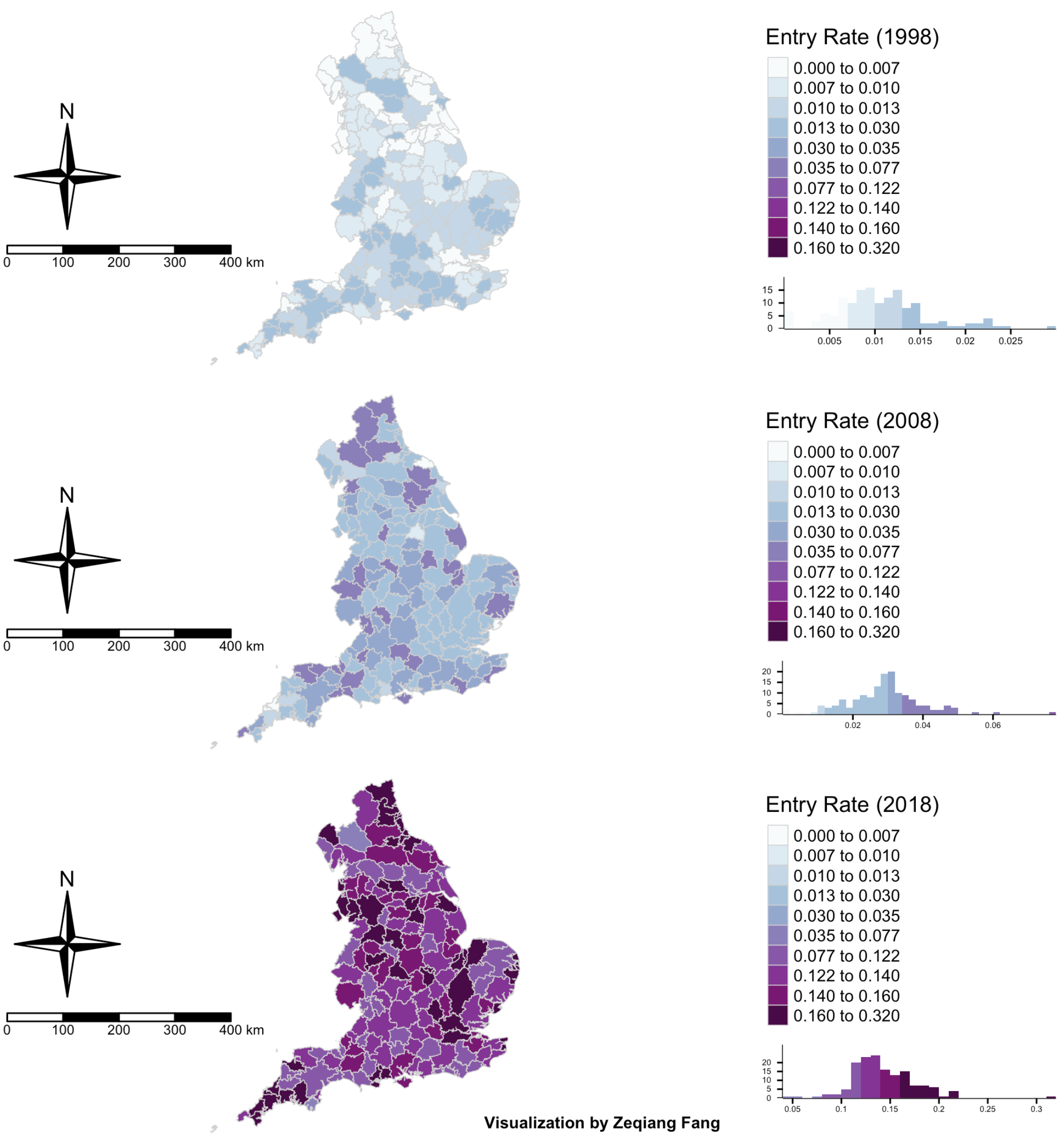

Figure 4.2: The Distribution of the Firms Entry Rate of the England Travel to Work Areas(TTWAs) from 1998 to 2018

power. The decline of the index means that the types of SIC codes corresponding to companies in a particular region have increased in a year or the number of companies corresponding to a large share of SIC codes has decreased. In other words, the types of technology industries corresponding to the unit area have increased, or the number of advantageous industries has decreased; you can see The concentration of technology industries out of England declined year by year from 1998 to 2018 (Table 4.1). In the same way, the diversity of the technology industry is also increasing, and the richness of the industry may also be related to the increase in the entry rate of enterprises.

From the perspective of temporal and spatial distribution, the technology industry was concentrated in northern England, northern and southwestern England before 2008. With the passage of time, even if the concentration of industries in some small areas scattered on the border of England is still high, the Herfindahl-Hirschman index in these places has gradually decreased to a level comparable to that of the Greater London area(Figure 4.3).

4.2 Regression Analysis

For 149 tech clusters(travel to work areas), the regression results can indicate the relationship among cluster dynamics, industry mix and firms density in different years(1998-2018) and locations (149 tech clusters). Among these three models, the model with outliers excluded and logarithmic processing of the dependent variable performed better. It can be seen that the Herfindahl-Hirschman index has a negative correlation with the tech cluster dynamic variable (entry rate) but the density has a positive effect.

Taking Model [3] as an example, every time this indicator decreases by one unit, the entry rate might increase by about $e^{1.37}$ unit times higher than the original one. The high confidence of this coefficient can explain 99% of all observation data. On the contrary, for the density of these tech clusters in Model [3], every time the density increases by one unit, the corresponding dependent variable entry rate will be about

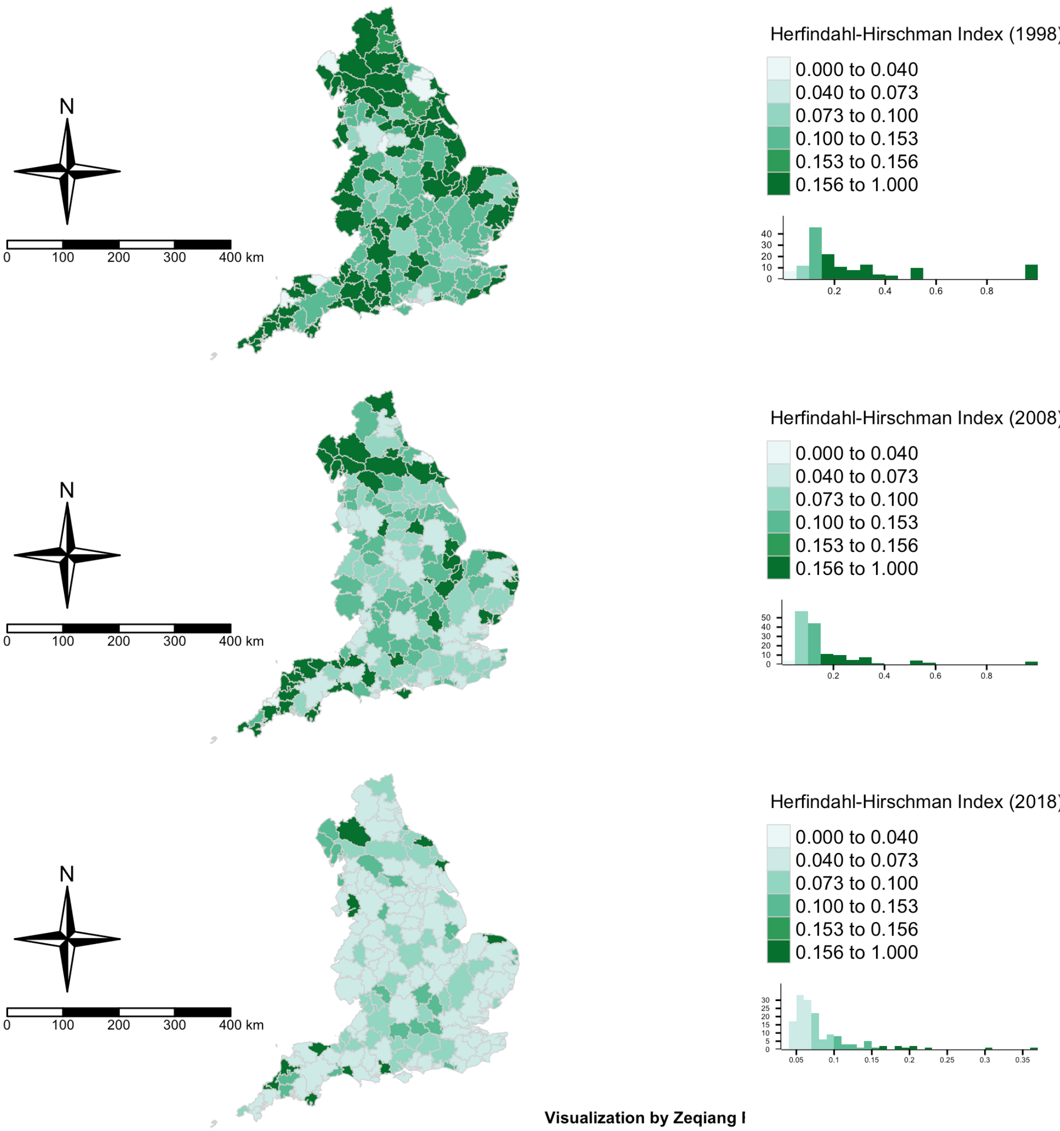

Figure 4.3: The Distribution of the Herfindahl-Hirschman Index of the England Travel to Work Areas(TTWAs) from 1998 to 2018

Table 4.2: Regression Results of Firms Entry Rate and Tech Clusters' Characters

Model	**[1]**	**[2]**	**[3]**
Method	OLS	OLS	OLS
Sample	all areas	all areas	exclude 2018 London
Dep. Variable	Entry Rate	Log(Entry Rate)	Log(Entry Rate)
HHI	-0.0281***	-1.332***	-1.3693***
	[0.002]	[0.045]	[0.044]
Density	0.0168***	0.3904***	0.504***
	[0.001]	[0.019]	[0.022]
Constant	0.0207*	-4.1138*	-4.0948*
	[0.003]	[0.068]	[0.066]
Observations	3089	3089	3088
R-squared	0.89	0.89	0.90

[a] All regressions include dummy variables for 149 travel to work areas(ttwa) in England and 20 years from 1998 to 2018. Figures in brackets are standard errors. ***, **, and * represent statistical significance at 1%, 5%, and 10%, respectively.
[b] HHI defined by Herfindahl-Hirschman index can indicate the industry concentration
[c] Source: tech firm data in England travel to work areas from 1998 to 2008 (n = 3089)

$e^{0.5}$ unit times more than the original one (Table 4.2). Although the HHI and Density values of the three models are not the same, the positive and negative effect of the relationship are similar. For example, for the industrial concentration variable HHI, all three models show a negative correlation (Table 4.2). And when HHI and Density change the same unit, the former seems to have a greater impact on the entry rate (it may increase the dependent variable entry rate by more than e times in both model 1 and model 2)

As mentioned in the methodology section(Brandt, 2005), Table 4.3 can illustrates the unconditional relation between entry rate and tech cluster characters. These two independent variables are controlled to be added in the regression model which can help to observe to what extent they can have an influence on the dynamics element(entry rate).

Table 4.3: Regression Results of Firms Entry Rate and Tech Clusters' Characters (Variables Controlled)

Model	**[1]**	**[2]**	**[3]**
Method	OLS	OLS	OLS
Sample	exclude 2018 London	exclude 2018 London	exclude 2018 London
Dep. Variable	Log(Entry Rate)	Log(Entry Rate)	Log(Entry Rate)
HHI	-1.1844**		-1.3693***
	[0.047]		[0.043]
Density		0.3779**	0.504***
		[0.025]	[0.022]
Constant	-4.1884*	-4.5143*	-4.0948*
	[0.072]	[0.075]	[0.066]
Observations	3088	3088	3088
R-squared	0.884	0.869	0.901

[a] All regressions include dummy variables for 149 travel to work areas(ttwa) in England and 20 years from 1998 to 2018. Figures in brackets are standard errors. ***, **, and * represent statistical significance at 1%, 5%, and 10%, respectively.
[b] HHI defined by Herfindahl-Hirschman index can indicate the industry concentration
[c] Source: tech firm data in England travel to work areas from 1998 to 2008 (n = 3089)

When the year and regional conditions in the model remain unchanged, the model [1] only adds Herfindahl-Hirschman index to consider the regression analysis, and it can be found that the coefficient is about -1.2 and the confidence level is greater than 95%. Similarly, the model [2] only considers adding firm density into the fitting, the coefficient of this independent variable is about 0.38 and the statistical significance is also less than 5%. However, when other conditions are the same, adding two variables into the model[3] at the same time, the absolute value of each coefficient increases, and the confidence level increases, and the overall goodness of fit is also improved which can be seen in table @ref(tab :tab-regression-control-var).

In terms of Herfindahl-Hirschman index, the entry rate might decrease when this indicator increase. Consistent with model[2], the entry rate increase with density (Table 4.3). The HHI decrease and density increase, leading to lower entry rates.

Table 4.4: Global Moran's Index of Firms Entry Rate in Three Time Periods

Year	**Moran's Index**	**p value**	**z-score**
1998	0.118	0.00	2.438
2008	0.062	0.08	1.348
2018	0.056	0.10	1.238

Table 4.5: Global Moran's Index of Firms Density in Three Time Periods

Year	**Moran's Index**	**p value**	**z-score**
1998	0.436	0.00	10.078
2008	0.325	0.00	7.545
2018	0.189	0.00	4.982

4.3 Spatial Autocorrelation

Through the global Moran's Index test of entry rate variable, it can be found that as time increases, the spatial cluster effect of entry rate attribute in England is getting weaker and weaker, from about 0.12 (1998) to 0.05 (2018).

The confidence of global Moran's I in 1998 is as high as 95%. But the statistical significance of 2008 is lower than 10%, which indicates weak interpretability of these two years(Table 4.4).

Contrary to the entry rate variable, the density attribute value of England has always maintained a relatively obvious spatial aggregation phenomenon, although the global Moran's Index of the density variable has been decreasing year by year. This variable has higher confidence in global Moran's I statistics in the three-year statistics, thus its interpretability is higher than that for mentioned entry rate (Table 4.5).

4.3.1 LISA Analysis of Firms Dynamics

Local indicators of spatial association(LISA) according to the table 4.6, the z-score with its significance can be found.

The visualisation result of the local Moran's index's z-score of Entry rate is as similarly

Table 4.6: Correspondence of Z-score and P-value

Z-score	**p-value**	**confidence**
less than -1.65 or more than 1.65	less than 0.10	0.90
less than -1.96 or more than 1.96	less than 0.05	0.95
less than -2.58 or more than 2.58	less than 0.01	0.99

indicated by the global Moran coefficient (Figure 4.2). With the growth of the year, the overall agglomeration effect of England has decreased. But, in this process, the agglomeration effect in the northern region was higher than that in other regions. However, the travel to work areas in the southwest corner changed significantly from the early cluster effect in 1998 to the corresponding outliers such as low-high, high-low areas in 2018 (Figure 4.4). In addition, from Figure 4.3, the higher industrial concentration in the north also makes the phenomenon of dynamics agglomeration more obvious than that in other places in England area.

To be more specific, high-high clusters in 1998 are mainly concentrated in travel to work areas in northern areas such as Newcastle, southern areas such as Andover-Newbury, and the southwest of Exeter. The entry rate values around these places may be higher than the British average. The low-low clusters of entry rates are clustered in the northern North Yorkshire area. The surrounding areas of these places usually have relatively close entry rates and low entry rate values.

The high-high cluster has shifted over time (the north like Whitehaven and the southwest corner like Launceston areas with higher entry rates gradually formed). By 2018, high-high cluster only appeared in travel to work areas around Liverpool, and new low tech firms entry rate areas were formed in Workington and Penrith&Appleby (Figure 4.4)

Central regions of England such as London, Birmingham and other regions have not had obvious spatial clustering effects in the visualisation results. This may be due to the low statistical confidence of the local Moran's index among regions. Or it might be because the data in this study is aggregated before calculated. The related data may be missing after the HHI and entry rate are utilised, which might not perfectly represent the company information in some parts of England.

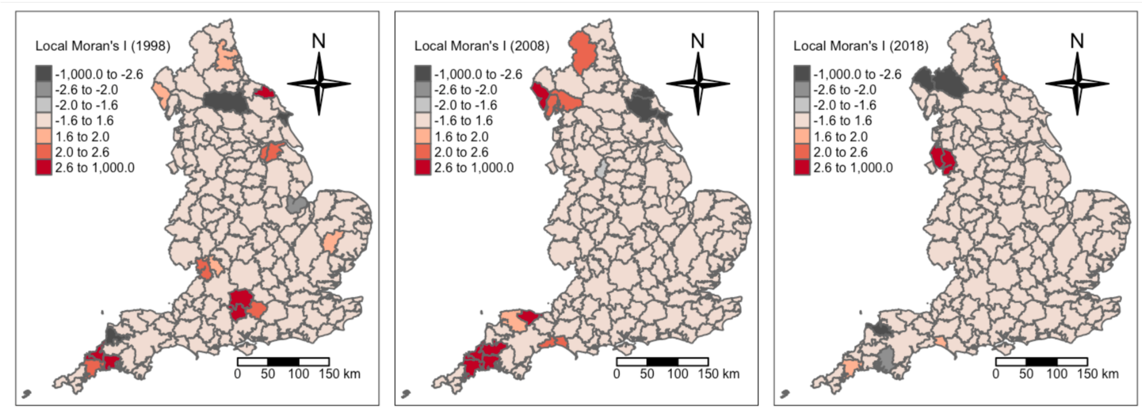

Figure 4.4: Local Moran's I Spatial Analysis for Entry Rate in England for Three Time Periods

4.3.2 Hot-Spot Analysis of Firms Dynamics

This research identified and mapped local clusters (hot and cold spots) by the level of entry rate for these three period.

On the whole, there is a distinct north-south difference in the distribution of cold and hot spots in the early stage(1998). With the development of England tech clusters, this difference is slowly dissipating and a reversal is formed eventually(Figure 4.5). More specifically, in 1998, there were some hot spots concentrated in the Swindon-Bristol-Reading urban belt in the south-central area and the southwest urban agglomeration of Exeter. Whilst, in the north, there were scattered cold spots represented by the Whitehaven area.

By 2008, most of the hot spots disappeared, and the places that were hot spots gradually became cold spots, such as the southwestern region. In 2018, some hot spots appeared in the northwestern region along with England. The original two hot spots in the southwest have experienced two-time nodes in 2008 and 2018, and they finally turned into cold spot clusters(Figure 4.5).

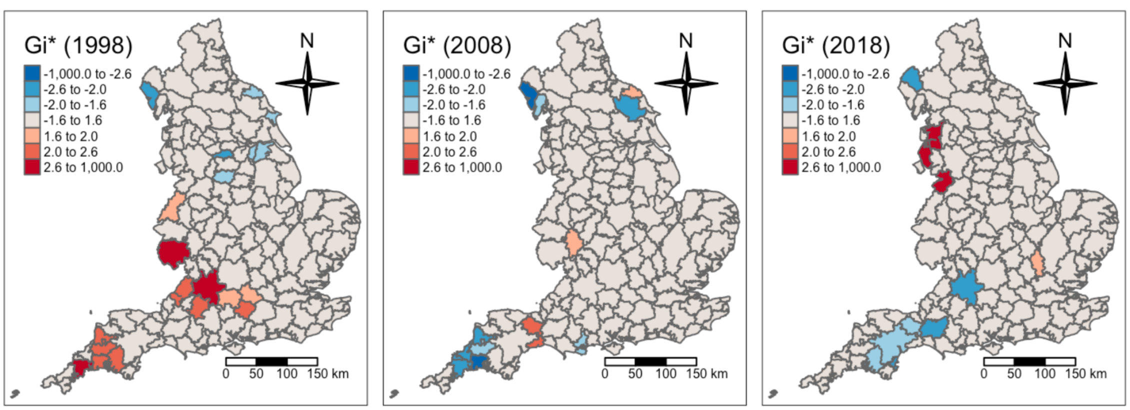

Figure 4.5: Local Gi* Score Statistics for Entry Rate in England for Three Time Periods

Chapter 5

Discussion

5.1 Reflection on Results

5.1.1 Spatio-Temporal Distribution

According to the trend shown in figure 4.1, in the 20 years since 1998, although the industry concentration in England fluctuated from 2002 to 2010, the overall average level has been reduced by about 70%. This means that after 2010, many new technology industry companies have joined the England region every year, resulting in a rising diversity of the technology industry. The British government (2016) mentioned in its official statistics report on industry concentration that Telecoms, which is more related to the technology industry, experienced a significant increase in indicators from 2006 to mid-2015 and then began to decline. Comparing the results in England, the overall concentration of the technology industry is in a downward trend. This phenomenon may affect the above-mentioned macro-industry trends.

Echoing the trend of increasing industry diversity is the rising trend of entry rate and company density. According to the results shown in figure 4.1 starting in 2016, the growth rate of the two indicators has suddenly increased compared to the indicator trend before 2008. This phenomenon may be due to the UK government(2017) plans to spend more than £1 billion to speed up the development and adoption of next-generation digital infrastructure in their policy paper. It can be seen that the prosperity of the industry, information and policies may promote each other.

From the perspective of geographical and temporal distribution, as an emerging technology cluster with a higher entry rate, northern England performed well after 2008 (above the average level). However, the central region, relying on Manchester and Birmingham and other regions, also has an above-average level in the medium term(2008). Combined with figure 4.3, the overall concentration of science and technology industries in the north is higher than that in the south. However, the industrial diversity in the London area has always been high, and there is no obvious trend of change. It is worth noting that the decline in industrial concentration in the central region has the most obvious effect. The corresponding growth in the entry rate of these regions is also obvious (Figure 4.2)

5.1.2 Firm Dynamics and Influence Factors

The regression results between the three variables found that in England, the more diverse the technology industry and the greater the density of technology companies, the greater the entry rate of regional technology companies. It is worth noting that the data of the London area in 2018 were excluded as outliers in the partial regression model. This is because the entry rate value in this year far exceeds the value in other years in other regions. This phenomenon may also be related to the rapid development of AI venture capital in the London area in 2018. For instance, between 2015 and 2017, the AI sector in London had a 200 percent growth in venture capital funding, and the city now has 13 universities offering AI-related degrees(Taylor,2018; Allott et al., 2018).

Meanwhile, the direct effect of firm density on firm dynamics seems to be more significant than the effect of industry concentration. Industry diversity often first acts on regional market performance and other factors, thereby indirectly affecting dynamics. Combining the large number of SMEs, mentioned in the literature section, tends to provide employment growth better than the small number of large companies

(Liu et al., 1999). Regions with large industry diversity may also have more small companies, which also explains the negative impact of high industry concentration on the increase of entry rate.

5.1.3 Spatio-Temporal Cluster Analysis for Two Indicators

Overall, the agglomeration effect of two indicators (firm dynamics and density) in England will reduce over time. However, the density index has a greater degree of reduction compared to the entry rate's cluster effect (Table 4.5). From the perspective of temporal and spatial trends, the north has gradually become a hot spot for the entry rate. More and more tech companies have joined the north and this phenomenon has local spatial autocorrelation. At the same time, these high entry rate regional clusters will be transferred to Liverpool and its surrounding areas in the later stage (2018). In the southwestern corner, the high entry rate hotspot from the beginning has gradually become a low entry rate cluster in the 10 years since 2008. How to implement policy encouragement and adjustments to make their technology industry rejuvenate may be a challenge to local authorities.

The LISA analysis of the firm density is used to supplement the exploratory description of the spatial model in this study. In figure 5.1, it might be a continuous pattern that London and its surrounding commuting areas have always had high-density clusters in these three periods although this spatial pattern may be weakened. Although there are also low-density clusters in the northern area, this might be related to the England railway line and transportation planning. The weakening effect of low-density clustering in this area is more obvious. These urban areas that are different from the London metropolitan area also have a certain absorption potential for new enterprises.

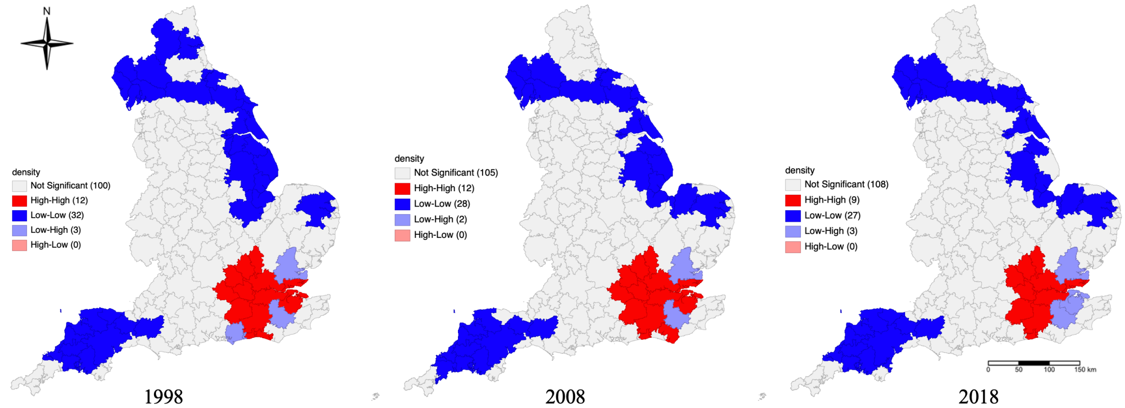

Figure 5.1: Local Moran's I Cluster Analysis for Firm Density in England for Three Time Periods

5.2 Limitations and Transferability

According to methodology, the method of identifying technology clusters mentioned in the section, the definition of the classification of technology companies comes from 2007. This division of the enterprise may not be the latest and most applicable method. Moreover, their information might be lagging because Office for National Statistics (2015) employed a comprehensive identification method, for example, refer to the code division part of *UK Standard Industrial Classification of Economic Activities* and the UK government official website *Science, technology, engineering and maths graduates in non- Industry definitions in STEM jobs* reports, etc. In addition, in the research of spatial pattern mining, there may be a lack of information in some areas. This will lead to low confidence in the local Moran's index detection, which cannot be discussed as a evidence. Finally, the study of hot-spot only calculates the significance standard and only considers univariate to participate in the calculation. The explanatory nature of spatial clustering analysis may not cover most of the scenes because of practical reasons. For example, the dynamics of enterprises are also affected by other factors in the space, such as policies, taxes, etc.

Studying other regions can also utilise the method of this study. For example, research on technology companies can start with company clusters and first quantify firm entry rate and industrial concentration. On this basis, the research can analyse the relationship between cluster characteristics and entry rate through regression models. To further study the spatial correlation of local areas, global and local Moran's index analysis methods and analysis of hot-spot to research the aggregation phenomenon of a certain index. This helps the government in a specific area to balance the development of each area to eliminate the Matthew effect of a certain industry, thereby realising the fair distribution of social production.

5.3 Recommendations

5.3.1 Corporate Decision-Making and Investment

Regardless of the influence of other UK regions except England area, the northeastern city belt might become a hot-spot for tech firms to enter. Although the southwest region was a cold-spot for that in 2018, the corresponding service infrastructure in the region is likely to be relatively complete because the region was a high entry rate cluster for technology companies to enter in 1998. Meanwhile, it can be found that commuters in a particular travel to work area often have a mature way of commuting from residential areas to commercial areas. This pattern can help us better understand the cluster phenomenon of technology companies. Moreover, when the home location is fixed, practitioners tend to choose work within a certain commuting area. In this case, it is more feasible for companies to hire high-end and stable labor in regions where market competition is less intense than to conduct the same corporate recruitment behaviour in other areas.

On the other hand, places with fierce market competition are likely to be highly attractive to talents, which will change the commuting decisions of individual families. Thus, company investors can also take into account the surrounding areas of London where there will be a greater tech firm entry rate in the future (Figure 4.2). Although this area is often treated as an outlier in the regression analysis results, London and its surrounding satellite cities may have very different corporate dynamics compared to other areas in England.

5.3.2 Government Policy and Strategy

Observing the overall situation, according to the experimental results, it could be found that the aggregation effect of the entry rate of enterprises shows a downward trend over time (4.4). From local analysis, the decline in the clustering effect in the northern area, especially the commuter areas in the northeast, is delayed compared to other areas. In terms of government decision-making, the implementation of policies in the northern and southern regions needs to be combined with the local real condition. For example, the government hopes to improve the average level of the number of technology companies in a large area in order to optimise the industrial structure. If the government of England aims to encourage local enterprises to move in or establish themselves, the investment priority can be considered in the cold-spot cluster area. Because increasing the enterprise entry rate in a certain area of the cold-spot cluster can eliminate the generation of low-low clusters to a certain extent in the spatial pattern. This is conducive to guiding the industrial balance between regions, thus weakening the Matthew effect.

From the perspective of the direct role of the government, such as the implementation of tax incentives for technology companies, and regional rental incentives to increase the entry rate of local enterprises. In terms of indirect action, the government can increase the entry rate of local enterprises by guiding the diversified development of regional science and tech industries and appropriately increasing the density of regional science and technology enterprises. According to the experimental results of (Table 4.2), the entry rate of enterprises in a certain area in a certain period of time is positively affected by the density of regional enterprises and the concentration of enterprises (Herfindahl-Hirschman Index).

Chapter 6

Conclusion

Aiming to make contributions to the research on the dynamic characteristics of tech clusters and industrial concentration, this article not only analyses the distribution of tech clusters in the England area but also conduct regression analysis and spatio-temporal pattern mining. In addition, the relationship between the entry rate of tech companies in England and the different indicators of the clusters (including firms density and the concentration of tech industries) is also quantitatively explained based on the results of the regression model and the spatial autocorrelation test.

In a nutshell, under the current research framework, whether it is the entire area of England or the local tech clusters, the experimental results in the past 20 years (1998-2018) have shown that industrial concentration and company density have a strong relationship with the entry rate of cluster technology companies. For example, the higher the industrial diversity and the greater the density of local technology companies, the higher the entry rate of technology companies in the region. In addition, the visualisation of spatio-temporal patterns also found the shift of cold and hot spots. For instance, compared with England as a whole, the cluster effect of the entry rate in some areas in the north has declined more slowly, while the southwest has changed from a cluster with a high entry rate to a cluster with a low entry rate. These findings can help government planning decision-makers to optimise the investment distribution of science and technology industries in England, thus realising the fair development of resource allocation among the England science and technology industries in the future.

Although the research framework has some limitations, the regression model and spatial autocorrelation analysis methods have good reproducibility in explaining the areas whose administrative divisions are similar to the UK. For the expansion of research directions, this research can refer to Clementi's (2016) analysis framework to investigate the dynamics of a cross-section of enterprises and their implications for aggregate dynamics. In addition, with the development of the technology industry and the trend of globalisation, the government's collection and archiving of more dimensional data of these companies would help to employ quantitative methods to explore the survival mode of tech companies. In addition to a more multi-dimensional quantitative analysis, this research can also focus on studying the impact of corporates' movement in the future, especially the change of the same company address and other corporate indicators such as the size of employees and assets before and after the relocation.

References

Alcacer J. and Chung W., 2007. Location strategies and knowledge spillovers, Management Science 53, 760–776.

Anselin, L., 2010. Local Indicators of Spatial Association-LISA. Geographical Analysis, 27(2), pp.93-115.

Agarwal, R. and Audretsch, D., 2001. Does Entry Size Matter? The Impact of the Life Cycle and Technology on Firm Survival. Journal of Industrial Economics, 49(1), pp.21-43.

Allott, S., Kingston, J., Endsor, M., Miller, M., Goldstaub, T., Sachania, M. and Janvrin, E., 2021. London: The AI Growth Capital of Europe. [online] Mayor of London. Available at: https://www.london.gov.uk/sites/default/files/london_theaigrowthcapitalofeurope.pdf [Accessed 23 August 2021].

Brandt, N., 2005. Business dynamics and policies. OECD Economic Studies, 2004(1), pp.9-36.

Burgess, G., Monk, S., Morrison, N. and Udagawa, C., 2014. Economic Analysis of the Wisbech Travel to Work Area: Main Report: March 2014.

Baptista, R. and Swann, G., 1999. A comparison of clustering dynamics in the US and UK computer industries. Journal of Evolutionary Economics, 9(3), pp.373-399.

Boschma, R., 2015. Do spinoff dynamics or agglomeration externalities drive industry clustering? A reappraisal of Steven Klepper's work. Industrial and Corporate Change, 24(4), pp.859-873.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D. and Roux, S. , 2012,The Productivity Advantages of Large Cities: Distinguishing AgglomerationFrom Firm Selection. Econometrica, 80: 2543-2594.https://doi.org/10.3982/ECTA84427

Clementi, G.L. and Palazzo, B., 2016. Entry, exit, firm dynamics, and aggregate fluctuations. American Economic Journal: Macroeconomics, 8(3), pp.1-41.

Carlsson B. ,1987 Reflections on industrial dynamics: the challenges ahead, International Journal of Industrial Organization 5, 133–148.

ESRI., 2021. How Hot Spot Analysis (Getis-Ord Gi*) works. [online] Available at: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm [Accessed 21 August 2021].

Ellison, G., Glaeser, E., Kerr, W. , 2010. “What causes industry agglomeration? Evidence from coagglomeration patterns.” American Economic Review, 100 (3), 1195-213.

Frenken, K., Cefis, E., and Stam, E., 2015. Industrial Dynamicsand Clusters: A Survey. Regional Studies, 49(1), 10-27.doi:10.1080/00343404.2014.904505

Fariñas, J. and Moreno, L., 2000. Review of Industrial Organization, 17(3), pp.249-265.

Feldman, M., 2003. “The locational dynamics of the US biotech industry: knowledge externalities and the anchor hypothesis.” Industry and Innovation 10 (3), 311-29

Fingleton, B., Camargo Igliori, D. and Moore, B., 2004. Employment Growth of Small High-technology Firms and the Role of Horizontal Clustering: Evidence from Computing Services and R&D in Great Britain, 1991-2000. Urban Studies, 41(4), pp.773-799.

Fritsch M., 2011. The effect of new business formation on regional development – empirical evidence, interpretation, and avenues for the further research, in FRITSCH M. (Ed.) Handbook of Research on Entrepreneurship and Regional Development – National and Regional Perspectives, pp. 58–106. Edward Elgar, Cheltenham

Furman, J. L., M. E. Porter, S. Stern., 2002. The determinants of national innovative capacity. Res. Policy. 31(6) 899–933.

Getis, A. and Ord, J., 2010. The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis, 24(3), pp.189-206.

Hall, B.L., Hsiao, E.Y., Majercik, S., Hirbe, M. and Hamilton, B.H., 2009. The impact of surgeon specialization on patient mortality: examination of a continuous Herfindahl-Hirschman index. Annals of surgery, 249(5), pp.708-716.

Hause, J. and Du Rietz, G., 1984. Entry, Industry Growth, and the Microdynamics of Industry Supply. Journal of Political Economy, 92(4), pp.733-757.

Kerr, William R., and Frederic Robert-Nicoud., 2020. Tech Clusters. Journal of Economic Perspectives, 34 (3): 50-76.https://www.aeaweb.org/articles?id=10.1257/jep.34.3.50

Kwoka, J., 1977. Large Firm Dominance and Price-Cost Margins in Manufacturing Industries. Southern Economic Journal, 44(1), p.183.

Krafft, J., 2004. Entry, exit and knowledge: evidence from a cluster in the info-communications industry. Research policy, 33(10), pp.1687-1706.

Kumari, M., Sarma, K. and Sharma, R., 2019. Using Moran's I and GIS to study the spatial pattern of land surface temperature in relation to land use/cover around a thermal power plant in Singrauli district, Madhya Pradesh, India. Remote Sensing Applications: Society and Environment, 15, p.100239.

Kogut, B., S. J. Chang., 1991. Technological capabilities and Japanese foreign direct investment in the United States. Rev. Econom. Statist. 73(3) 401–413.

Kwon, O., 2020. A study on how startups approach sustainable development through intellectual property. Sustainable Development, 28(4), pp.613-625.

Liston-Heyes, C. and Pilkington, A., 2004. Inventive concentration in the production of green technology: A comparative analysis of fuel cell patents. Science and Public Policy, 31(1), pp.15-25

Lu, C., Qiao, J. and Chang, J., 2017. Herfindahl–Hirschman Index based performance analysis on the convergence development. Cluster Computing, 20(1), pp.121-129.

Liu, J., Tsou, M. and Hammitt, J., 1999. Do small plants grow faster? Evidence from the Taiwan electronics industry. Economics Letters, 65(1), pp.121-129.

Li, L., Jiang, Z., Duan, N., Dong, W., Hu, K. and Sun, W., 2011. Police Patrol service optimization based on the spatial pattern of hotspots. Proceedings of 2011 IEEE International Conference on Service Operations, Logistics and Informatics,.

Mateos-Garcia, J. and Bakhshi, H., 2016. The geography of creativity in the UK. London: Nesta.

Malerba F., 2006. Innovation and the evolution of industries. Journal of Evolutionary Economics 16, 3–23.

Nathan, M. and Rosso, A., 2015. Mapping digital businesses with big data: Some early findings from the UK. Research Policy, 44(9), pp.1714-1733.

Open Corporates, 2018. The core company data from OpenCorporates master company database, electronic dataset, OpenCorporates Data Dictionary, viewed 12 June 2021, https://opencorporates.com/[Accessed 18 August 2021].

Office for National Statistics, 2015. Identifying Science and Technology Businesses in Official Statistics. [online] London, UK: Office for National Statistics, pp.10-14. Available at: https://webarchive.nationalarchives.gov.uk/20160105170025/http://www.ons.gov.uk/ons/site-information/using-the-website/rss-news-feeds/index.html [Accessed 28 July 2021].

Ozkul, B., 2014. Changing home-to-work travel in England and Wales. Regional Studies, Regional Science, 1(1), pp.32-39.

Prothero, R., 2021. Travel to work area analysis in Great Britain. [online] Office for National Statistics. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/traveltoworkareaanalysisingreatbritain/2016 [Accessed 8 August 2021].

Porter, M., 1998. “Clusters and the new economics of competition.” Harvard Business Review.

Porter, M E, and S Stern., 1999. The New Challenge to America's Prosperity: Findings from the Innovation Index (Council on Competitiveness, Washington DC).

Stankov, U. and Dragićević, V., 2015. Changes in the spatial pattern of net earnings: Evidence from Serbia. Acta Oeconomica, 65(3), pp.351-365.

Shaver, J. M., F. Flyer. 2000. Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. Strategic Management J. 21(12) 1175–1193.

Titheridge, H. and Hall, P., 2006. Changing travel to work patterns in South East England. Journal of Transport Geography, 14(1), pp.60-75.

Tech Nation, 2021. The Future UK Tech Built. [online] Tech Nation. Available at: https://technation.io/report2021/#methodology [Accessed 10 August 2021].

Tremblay, V., Iwasaki, N. and Tremblay, C., 2005. The Dynamics of Industry Concentration for U.S. Micro and Macro Brewers. Review of Industrial Organization, 26(3), pp.307-324.

Taylor, C., 2018 London is fast becoming a major hub for A.I. development. [online] CNBC. Available at: https://www.cnbc.com/2018/11/15/london-is-fast-becoming-a-major-hub-for-ai-development.html [Accessed 10 August 2021].

UK GOV, 2016. BIS analysis of key sectors. UK business sectors: indicators of concentration. [online] Department for Business, Innovation & Skills, UK GOV. Available at: https://www.gov.uk/government/statistics/uk-business-sectors-indicators-of-concentration [Accessed 23 August 2021].

UK GOV, 2021. UK Digital Strategy 2017. [online] UK GOV. Available at: https://www.gov.uk/government/publications/uk-digital-strategy/uk-digital-strategy [Accessed 23 August 2021].

Appendix A Data Source

Table 6.1: Research Data Source and Link

Name	Link
Dissertation Book Live Version	https://zeqiang.fun/CASA0012-Dissertation/bookdown/html/
Data Source Repository	https://github.com/fang-zeqiang/Master-Dissertation/tree/main/Dat
UK Registered Firms	https://opencorporates.com/info/our-data/
Science and Technology Classification	https://github.com/fang-zeqiang/Master-Dissertation/blob/main/Dat
Travel to Work Area Boundaries	https://github.com/fang-zeqiang/Master-Dissertation/tree/main/Dat
Regression Data	https://github.com/fang-zeqiang/Master-Dissertation/blob/main/Dat
TTWAs with Three Indicators in 1998	https://github.com/fang-zeqiang/Master-Dissertation/blob/main/Dat
TTWAs with Three Indicators in 2008	https://github.com/fang-zeqiang/Master-Dissertation/blob/main/Dat
TTWAs with Three Indicators in 2018	https://github.com/fang-zeqiang/Master-Dissertation/blob/main/Dat

Appendix B Research Log

Table 6.2: Research Log and Reproducible Analysis Process

Name	**Link**
Exploratory Data Analysis	https://zeqiang.fun/CASA0012-Dissertation/jupyterlab/EDA.html
Descriptive Analysis for Dynamics	https://zeqiang.fun/CASA0012-Dissertation/jupyterlab/Dynamics-Re
Regression Analysis	https://zeqiang.fun/CASA0012-Dissertation/jupyterlab/Regression-R
Spatial Pattern Research	https://zeqiang.fun/CASA0012-Dissertation/jupyterlab/Spatial-Resea
Map Making in R	https://rpubs.com/fangzq/UK_ttwa_tmap_making