Analyzing Urban Scaling Laws for economic Industries: Case Study of England and Wales, United Kingdom

[1]Jyothi Gupta
[1]Architect, Geospatial Analyst, Data Scientist, Smart cities and Urban Analytics
[1]The Bartlett Centre of Advanced Spatial Analysis (UCL CASA)
[1]University college London, WC1E 6BT, United Kingdom

Abstract - **This paper abstract is constructed to study the urban scaling laws in the UK context stating the problem, the research and the results in summarized way. The problem of this study is to identify the type of slope that we expect from the data. We shall analyse the scaling behaviour of (i) number of Knowledge based industries and Scientific services firms against city sizes, and (ii) the difference of these two industries along with their changes between the time frame of 2013 and 2018. Furthermore, I will investigate the difference in their behaviour and emphasize on the previous study whether they say anything about these both industries. This research is using the data from the company house and population census of the UK. This paper will answer three research questions. The first question is What are scaling laws and why is it used in this study. Secondly, we shall explain the knowledge-based industries in the UK and why it is selected for this research. Finally, what are the different types of slope we expect to find from the data. The scaling laws for England and Wales are explored for its linear pattern using this empirical analysis. The results for Knowledge based industry data scales sub-linearly for most types of industries whereas for scientific activity industry it shows the super-linear relationship. In other words, Scientific industry is positively scaling that is there been larger quantities of industry per capita in bigger cities. While concluding, we discuss the importance of using urban scaling law in England and Wales, variation seen in results from this empirical theory defined previously. This paper investigates the validation of the scaling model for city sizes.**

keywords - **Cities, scaling law, Exponent, agglomeration economics, linear**

I. INTRODUCTION

Background and context

Cities are complex system displaying patterns of inequality through agglomeration of space.(Batty, 2008) People, income, land use, other services and industries tend to agglomerate in different forms of cities (Gill and Goh, 2010). Besides complex system, a theory of how cities evolve, connecting with city economics and movement of human behaviour are developed as part of the city geometry.

The theme of this paper is what kind of slope we expect to predict from the knowledge-based industries of UK as per city sizes. Here we focus on to explain the urban scaling laws and why is it been used for cities, within agglomeration economics. We use this theory explained by Santa-Fe Institute particularly Geoff West and Luis Bettencourt. Also, using the data from UK industries, we analyse two types of industry that is the knowledge-based industry and Leisure industry for consecutive year 2013 and 2018. We are testing a theory in the cities of UK: that is that when cities get bigger, they perform more graciously quantitatively.

According to the West-Bettencourt study we have a theory. This empirical perception presented from this theory will be analysed. In this report we will test the theory whether it is a case or not for the UK and reason for selection. So, first with help of the literature will create a strong framework to start the hypothesis. Using company house data in UK, we will consider knowledge-based industries and Leisure industries and their difference will be measured. Also, will explicitly outline why am I considering these two types of industries from company house data of the UK. Indeed, scaling laws is studied in Physics, Biology and other fields which explains the relationship of the two physical quantities in terms of powers statistically.

Motivation of the study

This paper is to research on the urban scaling laws in the UK with respect to specific economic sectors and see whether there are any correlations regarding the performance of the cities and inequality indicators. For the economic sectors I will be looking at SIC codes in Companies data for two different points on time, 2013 and 2018, for specific sectors which will be knowledge-based industry and Retail industry. Moreover, we study the slope of the model and predict its behaviour.

I started with the papers on scaling laws from the class lecture Elsa gave on Urban System theory module in UCL CASA program. The main focus on literature will be on Bettencourt PNAS 2007, Elsa paper on Interface, Cottineau's paper, and Leitao's paper. It is aimed to find out the scaling relationships from the data and build a prediction in order to test the hypothesis with the city size of the UK. Indeed, it is very interested to investigate the spatial features of the economics of scales in UK supported by the Marshall's theories of agglomerations. It is indeed the case that many great cities form the past still stand today, but many other have appeared and disappeared as well.

Research question

The goal of the study is to answer the following three questions:

1. What are scaling laws and why is it used?
2. What is knowledge-based industries in the UK and why is it considered?
3. What is the type of slope we expect from the data? Explain the scaling behaviours analysis:
   a. The number of Knowledge based industries, Scientific services firms and leisure industry against city sizes.
   b. The difference and changes regarding time 2013 and 2018.

This paper is to research on the urban scaling laws in the UK with respect to specific economic sectors and see whether there are any correlations regarding the performance of the cities and inequality indicators. For the economic sectors I will be looking at SIC codes in Companies data for two different points on time, 2013 and 2018, for specific sectors which will be knowledge-based industry and Retail industry. Moreover, we study the slope of the model and predict its behaviour.

Research expectation

With the methods developed from Bettencourt and colleagues at the SFI, we expect the number of knowledge-based industries increases returns to scale by city size, showing that the local economics within particular industries are developing superlinear effects. This implies that the social economics of the whole city in the UK is showing superlinear scaling. However, we believe that some specific knowledge-based industries and business add to the overall superlinearity. We need to find that it's just the size of the city is important to plan new businesses and other economic activities that will be beneficial to understand the urban scaling behaviours.

In spite of the complications in this study, we test the cities with its distribution of industries, with the observation of the scaling exponents. With vast theory and detailed literature,we try to understand how city growth can be determined by the effects of interacting entities of urban indicators.(Pumain et al., 2006). Further, the city size in the urban scaling theory is the framework for measuring the fluctuation based on the fractal property of the city.

It is important to note whether we expect the same type of behaviour for the both the industries or not. Just to explain why we are considering two sets of industries. As per literature, we realise majorly most of the urban indicators scale super-linearly. We shall explain why we would expect the knowledge-based industry to be super linear or linear. It will expect how the data will fit in the urban scaling laws. Indeed, with new digital technologies, manufacturing sectors and other KBI, the calculated size of city increases with growth in number of firms that is to scale superlinearly. Certainly, the retail is dependent on the number of people, income and employment as major growth in the city. According to literature majority of the urban indicators are superlinear, let's see what our results from company house data will agree with this empirical theory.

Research objective

To better understand an agglomeration economics that defines the city's growth system. It is important to understand the urban scaling empirical theory that measure the scaling behaviour of city size. We produce analysis of urban scaling in England and Wales, by aggregating the knowledge-based industries along with population census data. We then measure the order of magnitude of scaling exponents to identify the type of slope in the data. While concluding we investigate the scaling variation with the time frame of 2013 and 2018 to link the economic scale of cities with their city sizes. Additionally, we also explain why the knowledge-based industries are observed to scale superlinearly as compared with Leisure industry from the company house data in the UK.

Paper Outline

The outline of this article is in the following arrangement:

Section two reviews the literature of urban scaling law and growth in the city complex system for recent years.

Section three introduces the data of England and Wales, data aggregating approaches and scaling methods.

Section four analyses the scaling (log-log OLS regression) and visualisation patterns for knowledge-based industry and leisure company house data for the two-time frame for year 2013 and 2018 and explores the scaling exponent results.

Section five broadens the discussion on the scaling exponent results and examines the influence of the different slope and limitation and significance of this study. Chapter six leads to explain the conclusions and future work of the scaling relationship for city sizes.

II. LITERATURE STUDY

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts. True Type 1 or Open Type fonts are required. Please embed all fonts, in particular symbol fonts, as well, for math, etc.

The work is split into two literatures; one is discussed about scaling laws empirical method and the second part researches about the agglomeration economics.

Growth in cities

Many people are living in cities. The cities are growing mainly in economics development which is primary to changes in land use, social organisation and patterns of human behaviour.(Bettencourt et al., 2007.) These changes are not equal impacting the global environment. Consequently, to analyse and predict these changes in social organisation, land use and human behaviour there is an empirical theory presented. Also, characters of cities are used for scaling exponents presented by power-law equation with population size. In his paper, the main topic was the nature of urban morphology, city size, urban allometry and additionally the qualitative changes occurring in the economics of cities as they agglomerate. This key subject about the world wherein every citizen prefers to stay in cities with many advantages and aligned disadvantages.

This paper is most important in this report particularly Geoff West and Luis Bettencourt work on this empirical theory. They have defined and studied the US data with 357 cities using Metropolitan statistical areas (MSA) data. As per the empirical method, the income grows more than proportionately beside the size of the city. So, the measure of city size at time, power law scaling is formularised below;

Equation 1. Scaling Law equation

$$Y = (m\, x^{\beta})$$

Y – Measure of social activity (wealth, pollution, energy, infrastructure)

x – Population size

m – constant or intercept

β – slope on the log-log plot

$Y = (m\, x^{\beta} \sim m\, x^{1.10})$, here Y is income and x is population size, then Beta is more than 1. This is called super-linear scaling $\beta > 1$.

The findings in West-Bettencourt's study signify about scaling of the urban system organisation. Cities are certainly unique in their form and magnitude as predicted with the help of scaling laws. In power scaling law, Bettencourt defines in table 1 that scaling exponents β are divided into three classes;

Linear Relationship where β = 1 related with needs of people (for example household, occupation, social collaboration) which are one.

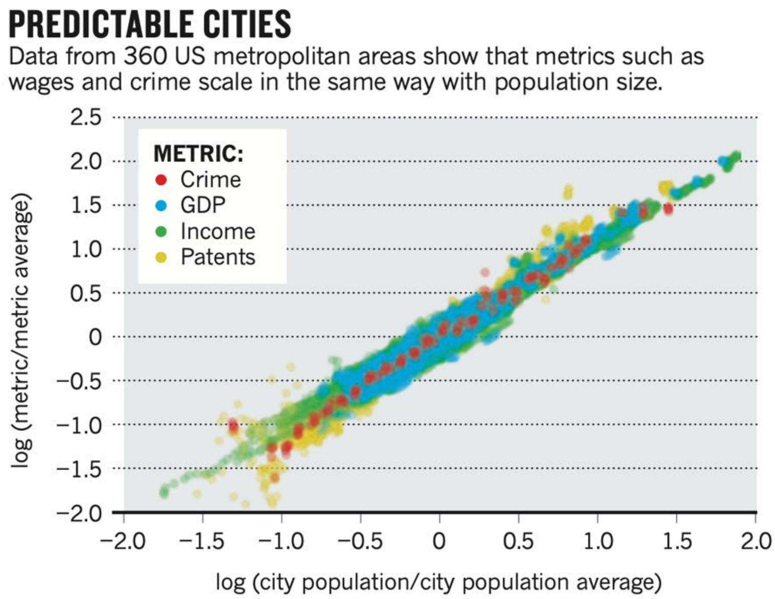
Sub-Linear Relationship where β < 1 related with material measures (like Infrastructure, economics, costs) which is more lesser than one.

Super-linear relationship where β > 1 related with increasing in population size, information, wealth and other social nature of cities: greater than one.

Table 1. Different sets of Scaling Exponents

Relation of scaling exponents	Beta related to One
Linear Relationship	$\beta = 1$
Sub-Linear Relationship	$\beta < 1$
Super-linear relationship	$\beta > 1$

Bettencourt from SFI summarised that with cities sustaining their growth, there is rise in innovation in cities with human behaviour scaling super-linearly. The process of data gathering, aggregation and comparison aim to explain the development in cities. Indeed, this research develops empirical theory about cities that analyse the quantitative measure in terms of social and economic activities like innovation, resource and consumption. Bettencourt's theory defined cites as complex system generating quantitative evaluation for creating a framework of understandings of urban



systems

Figure 1. Variations of urban scaling indicators with population

Let's take this example of US cities in figure 1 which shows the spread of data measuring each city as per its sizes. The metric with four urban indicators which is distributed based on their quantities in the urban centre of the developed cities. On average, with increasing in city size, the socio-economic quantities such as number of firms, GDP etc. increased by 15% expected linear growth. However, other damaging activities like traffic, diseases, crime for the city also increases by 15 % rule.(Bettencourt and West, 2010) This empirical observation in urban system clearly explains social organisation as observed in scaling exponents. This paper fits the data using OLS method and then used logarithmic binning to explore the exponents β.

Agglomeration economics

Agglomeration economics is an important topic for policymakers who is offering public policy proposal for the city.(Glaeser, 2010) Basically, the number of firms and citizens are located close to each other in cities and industrial clusters. Ultimately, this concept of agglomeration economics is cost saving method which relates to economies of scale and network effects. Cities form and grow facilitating the economics of agglomeration. Cottineau emphasised whether or not larger cities are more valuable and useful than smaller cities. The author measures the agglomeration economics for spatial and economic complexity of cities showing both quantitative and qualitative change.(Cottineau et al., 2018) With a detailed study of French cities, Cottineau observe that the big cities tend to concentrate the jobs rather than higher wages(Cottineau et al., 2018). Indeed, for income, the cities concentrate centrally to have more income per capita than wages per job. Henceforth, when analysing the agglomeration economics for French cities the location of house and job within the cities is important. So, the central hub of the city tends to spatially produce more income in cities with large size. On the other hand, empirical evidence produced Glaeser view for US cities is that larger cities are more productive because they attract the skilled workers. (Glaeser, 2010) These researchers direct that the statistical measurement for the nonlinearities could be a major focus for providing better measurements of agglomeration economics effects. Also, with this literature we can come to an opinion that the agglomeration economics applies different results in different location and policymakers. It's not always true that larger dense cities are more productive over time.

Heteroscedasticity

The growth of cities inspired early research in urban system theory, but this session will focus on the inequalities in the size of the city. (Sarkar, 2018) This paper explores the quantitative and qualitative effects of scaling to the geographic strengths of cities. Sarkar overall results explains the dissimilar fluctuations seen in the data. She creates scatter plot along with the city sizes showing the behaviour relating the scaling analysis test. We define heteroscedasticity in this paper which is about the un-similar variations predicted from the data in the results which are either super-linear, or there are

disproportionate agglomerate of the socio-economic quantities. The model attempts to determine the objective and scaling relationships between size of the city and inequalities in urban area. The empirical findings discussed that due to the concentration of high-income employment and high land cost for housing, we notice that the low-income crowd is excluded for not able to afford. However, this clause will lead to create a much-organised urban policy in the cities which create a balance in diversity, affordability and socio-economic income distributions. This study uses the data of Australia and USA to impact various scaling relationships and strategies on the performance of their cities. Indeed, Sarkar concluded that this asymmetric distribution of urban system needs to concentrated and main objective is related to agglomeration and city size.

Quite a few researchers have understood this empirical theory and the problem relating this scaling behaviours to optimise and create a framework. Leitao research also validates that the data with order of magnitude showing heteroscedasticity as the urban indicators are larger and fluctuates with the population size of the city. Nonetheless, this related assumptions, enhancing the cities inequalities are being compared with the quantitative analysis proposed by the linear model.

Well, there is a hypothesis test in inequalities and scaling of a city size. Traditionally, there is a distribution of data with urban economics models. The basics is about the income, housing rents, infrastructure transport cost which is totally come to the standard of living in the city. However, these observations are very uniform with its distance to the significant hub of the city centre and capturing the heteroscedasticity of the real cites in this scaling model. Further, we will like to explain the equation for both the exponential and power law distribution.

Building a city

Thanks to the recent researchers (for instance M.Batty, E.Arcaute, Cottineau, Bettencourt, West and others) for providing major contribution concerning city performance and size. This section we will summarise the paper Arcaute et al 2015 wherein the scaling laws was not used for constructing cities, quite the opposite, cities were constructed from the bottom-up approach focusing at population density. Indeed, this research helped to test the scaling laws in many different systems of cities. Although Marshal coined the urban agglomeration, it is indeed being important subject to study quantitative change in cities with its size. With the findings from the West-Bettencourt Research for US cities were not very appropriate for cities in the UK. However, the empirical analysis shows the size of the city with many urban indicators scale linearly using scaling laws for the census data from the UK.(Arcaute et al., 2014) Also, the sensitivity analysis for the scaling exponent β is well explained using the heat maps to represent the β showing the income, population size does not adhere the agglomeration economics of scale and deviate from the Bettencourt results. This superliner scaling exponent which was predicted for the income scale with population showed homogeneity for all different cities. Hence, this exponent result is sensitive to England and wales census data. When the result is been compared with Bettencourt US (MSA data) scaling exponent along with Arcaute E&W data, the discrepancy leads to a new indication that London is considered as outlier. Also, the dragon-kings theory referring Sornette's idea of city theory emerged to economic hubs which are not necessarily correlated to the size of the city. To conclude, the scaling laws which are used for measuring the performance of a city are not very successful for all the city location and defined size. This theory cannot be relied for predicting the behaviour in diversity and heterogeneities of cities.

Variations of their behaviours

To estimate and frame the scaling laws, the variation of scaling exponents β defining the urban system (Cottineau et al., 2015). Cottineau explains this variation of scaling with four steps. (i) Clustering (ii) Heatmaps (iii) Hierarchical clustering (iv) Extreme scaling. She summarises about city having similar attributes across its size. Hence, this helps to identify the cities from its behaviour when the population size is matched. However, there were deviations recorded in the results which the author explained with its attribute location. So, the variations are neither random nor standard for all the urban indicators under study. Also, the assumption in this paper for cities has been monocentric but could be further enhanced with the diversity of urban forms. Finally, the variation in scaling is completely dependent over time for economic and sociologic categories. Definitely, this measures the magnitudes for scaling exponents explored for the data in France predicts the value and covariation between population and urban indicators.

Further, we discuss the results of E&W cities (Arcaute et al., 2014) from Arcaute paper showing the scaling analysis for cities and its urban indicators.

Nonlinearity Scaling

This session will summarise the research from Leitao et al, 2016 from royal society.

The scaling formula as scripted earlier;

Equation 2. Power Scaling law equation

$$y = (mx^{\beta})$$

β = 1 Linear scaling β > 1 Super-linear scaling

β < 1 Sub-linear scaling β ≠ 1 Non-linear scaling

The author will test the exponent and CI and observe the nonlinear exponent and explain the fluctuation of the data. Also, will model the fact the city sizes are heavy tailed to normal distribution. (Leitão et al., 2016) Indeed, this literature shows comparisons of the residuals of the fit to the gaussian distribution predicted by the linear fit log-log model. This statistical framework for testing the scaling exponents regardless of linear or non-linear scaling provides a platform for residual analysis method. However, the findings of this study lead to rejection of many models (p value < 0.05) considering the significance level of data analysis and many other measurement reasons. The exception of the results where the log-log models in UK- income and train stations were compared to introduce better result. (Leitão et al., 2016) Nevertheless, the scaling laws in cities were analysed and compared with β parameter in the data and concluded that the variation of β across different models and methods proofs β fluctuation of greater and smaller to 1 and not equal to 1.This approach of the least-square fitting is proposed in this paper with help of pearson, gaussian and lognormal models and method of analysis leading to non-linear scaling. So, the city population depends on the hypothesis of these fluctuations of the scaling exponent β. Consequently, this result was not efficient to conclude the evidence of scaling to be non-linear as investigated scaling laws beyond cities. (Leitão et al., 2016)

III. DATA AND METHODS

For this study, we use three set of data for England and wales.

UK data

For this study, the company house data is collected for England and Wales for the time-period 2013 and 2018 from www.gov.uk which included the registration details with five-digit number, incorporation dates, address with postcode, company type and status. So, the data size of 19,00,000 (1.9 million/ 19 lakhs) rows and 21 columns. Standard Index classification – SIC provides a unique number which is further divided in hierarchal mode with their type of manufacturing, its dependent activity.

We choose two classifications of SIC numbers as presented in table 2. In this study the data used are for knowledge-based industry with type 5 classes and leisure industry with type 3 classes. The data are explored in the spreadsheet and analysed with its nominal, categorical, binary variables. The whole process of data prepossessing is carried out on the company house data in order to work on the scaling models. Later the data are transformed to the csv format and linked to R program to help create the spatial analysis and create basic visualisations of bars, pie charts and few interactive charts in Python language. This data are repeated for two consecutive years 2013 and 2018.

The data points in KBI with five categories, Leisure with three modules distributed both since year 2013 and 2018 is visualised in vertical barchart in figure 2. This grouping is observed with their activities in the urban built environment.

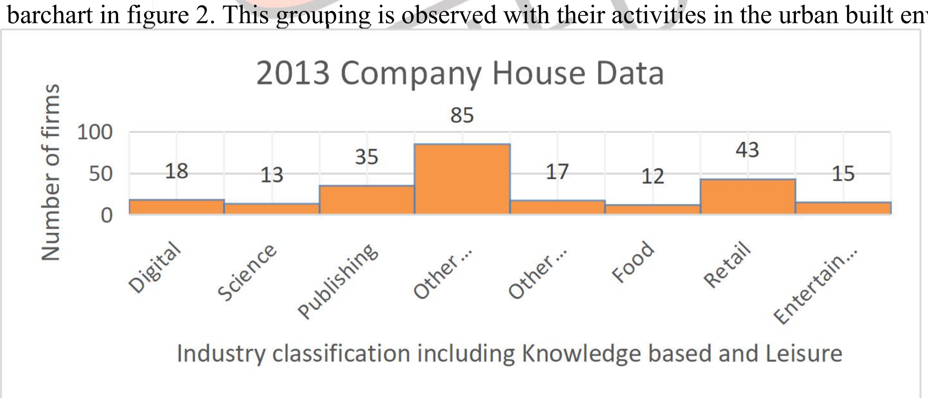


Figure 2 Histogram chart representing the 2013 company house data for both the industrial classifications.

Table 2.Company House Data for the UK for Knowledge based Industries

	Knowledge based industries SIC classes	Description
Type 1	Digital technologies	
1.	26110, 26120	Manufacture of computer, electronic and optical products
2.	26200,26400,26511,26512,26800	Manufacture of consumer electronics and Manufacture of instruments and appliances for measuring, testing and navigation; watches and clocks
3.	33130	Repair and installation of machinery and equipment
4.	58210,58290	Publishing of computer games and Other software publishing
5.	62011,62012,62020,62030,62090	Computer programming, consultancy and related activities
6.	63110, 63120	Information service activities
7.	95110	Repair of computers and peripheral equipment
Type 2	Life Sciences and healthcare	
1.	21100,21200	Manufacture of basic pharmaceutical products and pharmaceutical preparations
2.	26600,26701	Manufacture of computer, electronic and optical products
3.	32500	Manufacture of medical and dental instruments and supplies
4.	72110,75000	Research and experimental development on natural sciences and engineering and Veterinary activities
5.	86101,86102,86210,86220,86230 ,86900	Human Health and Social Work Activities
Type 3	Publishing and broadcasting	
1.	26301,26309,26702	Manufacture of communication equipment (other than telegraph and telephone apparatus and equipment) and Manufacture of photographic and cinematographic equipment
2.	58110,58120,58130,58141,58142 ,58190	Publishing of books, periodicals and other publishing activities
3.	59111,59112,59113,59120,59131 ,59132,59133, 59140,59200	Motion picture, video and television programme production, sound recording and music publishing activities
4.	60100,60200	Radio broadcasting; Television programming and broadcasting activities; Programming and broadcasting activities
5.	61100,61200, 61300,61900	Satellite telecommunications activities, Wired telecommunications activities
6.	63910,63990	Other information service activities such as Data processing, hosting and related activities; web portals
7.	73110,73120,73200	Market research and public opinion polling Advertising and market research
8.	74100,74201,74202, 74203,74209	Other professional, scientific and technical activities, Photographic activities

9.	95120	Repair of computers and communication equipment
Type 4	Other scientific activities of manufacturing	
1.	19201,19209	Other treatment of petroleum products (excluding mineral oil refining/petrochemicals manufacture)
2.	20110,20120,20130, 20140,20150,20160,20170,20200 , 20301,20302,20411, 20412,20420,20510,20520,20530 ,20590,20600	Manufacture of man-made fibres, Manufacture of other chemical products, Manufacture of paints, varnishes and similar coatings, printing ink and mastics
3.	25210, 25300,25400	Manufacture of tanks, reservoirs and containers of metal
4.	26513,26514,26520	Manufacture of instruments and appliances for measuring, testing and navigation; watches and clocks
5.	27110,27120,27200, 27310,27320,27330,27400,27510 ,27520,27900	manufacture of products that generate, distribute and use electrical power; manufacture of electrical lighting, signalling equipment and electric household appliances.
6.	28110,28120,28131,28132,28140 28150,28210, 28220,28230, 28240,28250,28290,28301,28302 28410,28490,28910, 28921,28922,28923,28930,28940 28960,28990	Manufacture of machinery and equipment like materials either mechanically or thermally or performs operations on materials (such as handling, spraying, weighing or packing
7.	29100,29201,29202,29203,29310 ,29320,	Manufacture of motor vehicles, trailers and semi-trailers
8.	30110,30120, 30200,30300,30400,30910,30920 ,30990	Manufacture of other transport equipment such as ship building and boat manufacturing, the manufacture of railway rolling stock and locomotives, air and spacecraft and the manufacture of parts thereof.
9.	32120,32401	Manufacture of games and toys; Manufacture of jewellery, bijouterie and related articles
10.	33120,33140,33150,33160,33170	Repair and installation of machinery and equipment
11.	95210,95220,95250	Repair of personal and household goods
Type 5	Other scientific activity services	
1.	51101,51102,51210,51220	Air transport - transport of passengers or freight by air or via space.
2.	71111,71112,71121,71122, 71129,71200	Architectural and engineering activities; technical testing and analysis
3.	72190,72200	Research and experimental development on social sciences and humanities, natural sciences and engineering
4.	74901,74902	Specialised design activities
5.	85410,85421, 85422	Post-secondary non-tertiary education, Tertiary education and Post-graduate level higher education

Table 3. Number of Industrial Classification from Company House data categorized in Knowledge based and Leisure Industry.

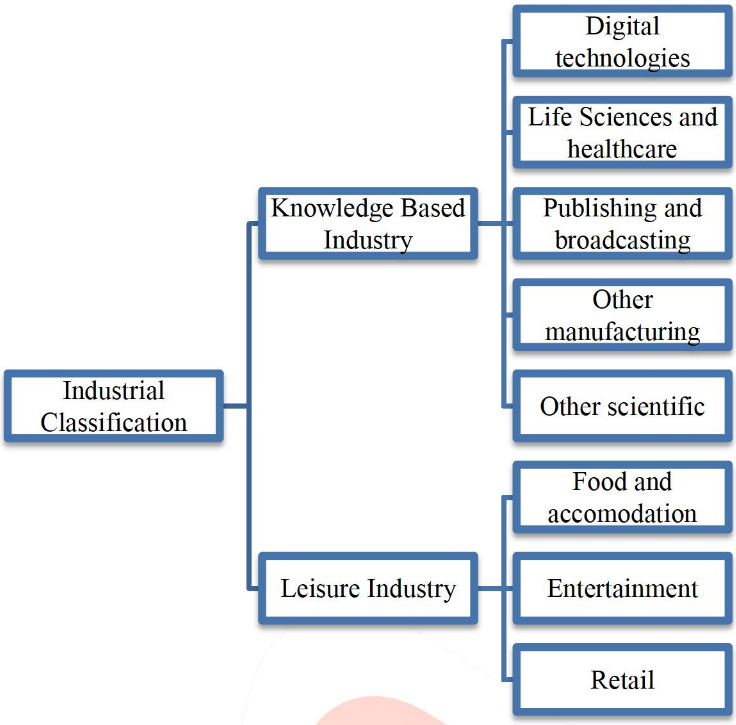


Table 4. Company House Data for the UK for Leisure Industries

	Leisure industries SIC classes	Description
Type 1	Food and accommodation	
1.	55100,55201,55202,55209,55300	Hotels and similar accommodation Holiday and other short-stay accommodation; Camping grounds, recreational vehicle parks and trailer parks
2.	56101,56102,56103, 56210,56290,56301,56302	Restaurants and mobile food service activities; Event catering and other food service activities; Beverage serving activities
Type 2	Entertainment	
1.	59140	Motion picture, video and television programme activities
2.	82301,82302	Organisation of conventions and trade shows
3.	90010,91011,91012,91020,91030, 91040,92000,93110, 93120,93130,93191,93199, 93020	Creative, arts and entertainment activities; Libraries, archives, museums and other cultural activities; Gambling and betting activities and Sports activities and amusement and recreation activities
Type 3	Retail except retail trade of motor vehicles	
1.	47190,47990,47290,47789,47782,47810, 47890,47820, 47781,47110,47791,47430,47300,47250, 47610,47240, 47530,47710,47410,47750,47540,47230, 47760,47721, 47210,47599,47650,47520,47741,47722, 47220,47749, 47421,47630,47591,47620,47799,47640, 47429,47510, 47260,47770,47910	Retail sale in non-specialised stores; variety of product lines in the same unit (non-specialised stores), such as supermarkets or department stores.

Postcode data

The next data we need in our study are the postcode data. This data are required to join with the company house data. This postcode CSV file is converted to shape file with help of ArcGIS command. Although the UK postcode is a large size data with its statistical points transformed using the spatial join. It is a polygon file which is clusters to the point file. The explanation was answered using R language and ArcGIS software. Additionally, the data were supposed to be in the wgs84 geographic coordinates system, so to measure distance. So, this data were transformed into the projected British National Grid coordinates system. While the coordinates are an important part for any spatial analysis, but if any uncertainty in the coordinates can destroy the complete GIS analysis. Further, with help of the spatial clip, it joins the points, export its new shape layer as seen in figure 3.

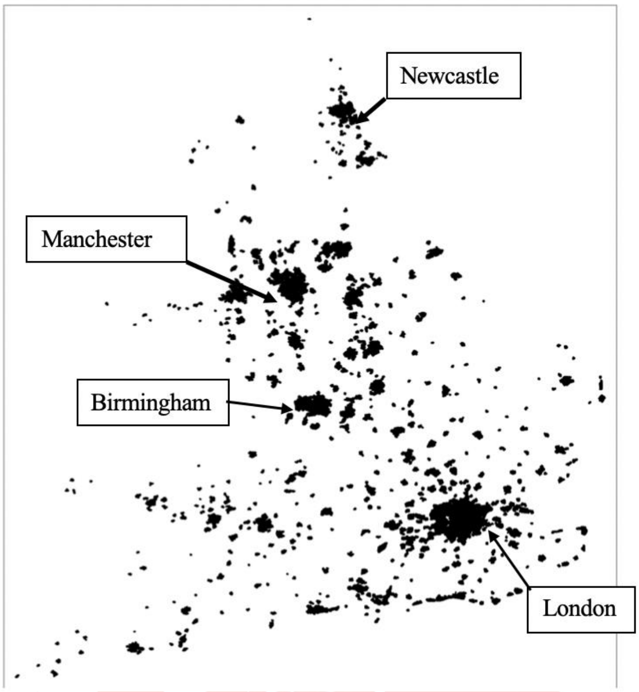


Figure 3.Spatial representation of England and Wales urban clusters of cities

Population census data

The Office of Nation Statistics helped to download the population of England and Wales for the year 2011 with a higher level of geographic area (ons.gov.uk).

1. Output Areas smallest with a population of 100.
2. Super-Output Area
    a. Lower Layer Super Output Area with a population of 1500.
    b. Middle Layer Super Output Area with a population of 7200.
3. Local Authority Districts

Population data was measured with the cities in England & Wales in the UK with the above three output areas in the UK. It is important to understand the data as we had to access it along with the other three data sets.

Join with UK clusters

The dataset was now in three buckets which had to be linked to create a complete spreadsheet. The main data processing tool used in this research is RStudio. We had to categorise the columns and rows related to the KBI and Population size of cities. We extracted the census data, postcode data with coordinates. By removing the duplicate rows, we get a clarity in the data for the UK regions only. Later we identify the names, MSOA, Latitudes and longitudes (coordinates system) of the data. This step lead to a delay process since the file size was large and taking time to merge the dataset. However, in some records, we find incomplete data or having NA values. Therefore, in order to avoid errors in the model, we cleaned the null NA values and finally obtained the figure 4 in ARC GIS software. The join was useful to identify the density and concentration of population of cities in the UK with help of the map. This spatial information was calculated with the count of the firms in the cities.

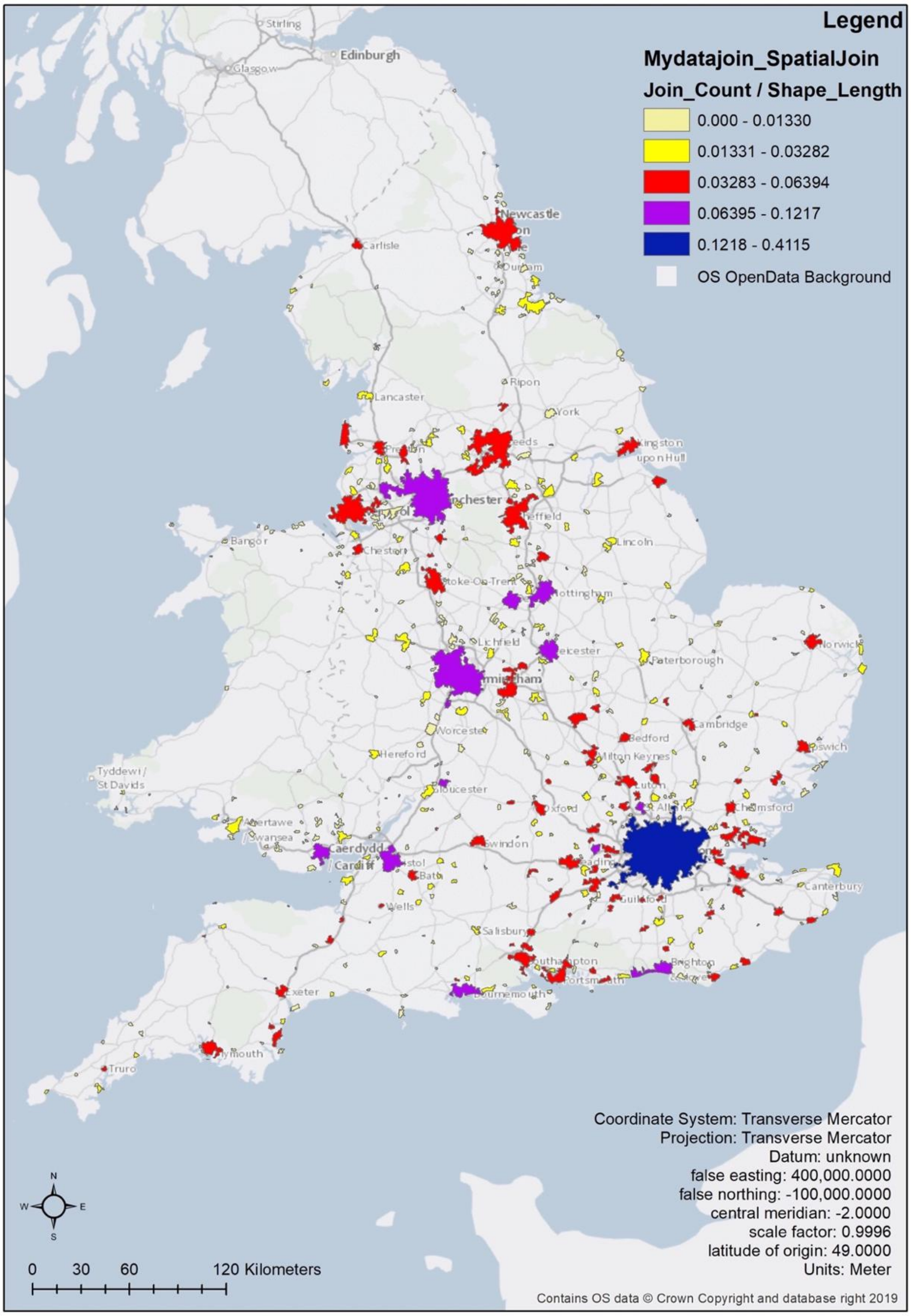


Figure 4.Cities defined in variation of sizes spatially with largest being London and smallest is York.

IV. EMPIRICAL RESULTS

Our observed results are planned into three classifications, following three questions we asked in the introduction. Section 4.1 answer the questions 'what are scaling law and why it is used' with verification of biology, physics and scientific methods. Section 4.2 explains the knowledge-based industries and reasons for consideration. Section 4.3 is the last question which examines the data to find the slope and its relation to the size of the city.

In this section, we study scaling of the aggregate number of firms in the log-additive model with city size. The UK data do expect a total scaling exponent to be higher than one, this value emerges to predict total number of industries in the larger cities.

We re-perform this analysis since 2013 and 2018 inputting the total population of cities and number of firms in all 8 categories of selected economic industries. Second, in each log-log regression model, we collate the scaling exponents to be in linear to superlinear range, calculating the increasing size of cities with industries' scaling behaviour. However, the result will be discussed along with the Bettencourt and other leaders in urban scaling theory predicting the correlation in our data.

What are scaling laws? And why it's used?

In biology dictionary, there is a variation seen organisms from microbes to whales spanning to more than 21 orders of magnitude in size. (West et al, 2000) This size affects the structures and metabolism to population dynamics (West, 1997). So, the quantities are metabolic rate, time scales and size of the organism. Indeed, the valid forms of life are expressed by power law relationship with exponents that are simple multiples of ¼. (1/4,3/4 ,3/8 ,…) (West, 1997)

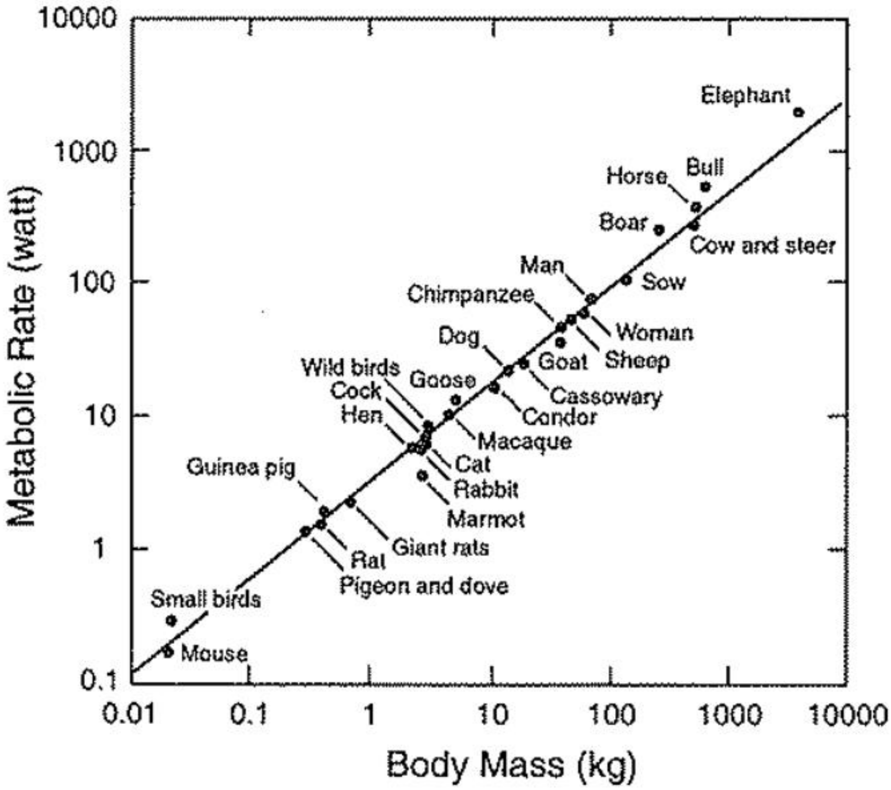


Figure 5.Biological representation of Metabolic rate of series of organisms created by West et al 1997 discovered by Kleiber

The allometric scaling relationship coined by Julian Huxley (West et al., 2002) focuses on aging and mortality of metabolism as in figure 5. This model describes structure of the biological network based on their hierarchical branching structure (West, 1997). The predicted scaling properties limits on fractal-like structure determining many scaling relationship explaining allometric scaling(West et al., 2002). Technically, the following equation with the variables defined which varies as per the data.

$$y = (mx^{\beta})$$
$$B = (B_0 M^{\beta})$$

y – biological variable - B
m – Normalisation coefficient – B_0
x – body mass - M
β - scaling exponent - β

In Physics, scaling laws is the quantitative relationship between length and area that scale with each other over a period of time. It can, however, be expressed both in Biology and Physics, but our study is for cities. Bettencourt highlighted by using this empirical theory to observe 'cities of different sizes do have different properties.' M.Batty presented scaling law with range of five orders of magnitude for examples bigger cities in order of millions 25m Tokyo, Mexico City and the smallest in order of 100's. Statistically scaling is reflected in fractal structures as we analyse the physical form of cities, hierarchy of road systems, etc.

In this paper, we built the scaling relationship with number of firms – KBI for England and Wales vs population census data for year 2013.The Line showing is the slope with scaling relation $y = (mx^{\beta})$, with β = 0.85; theorical prediction of the sub-linear slope. The observations are summarised for different KBI firms and Leisure industry which largely depends on the city size, usually measured with the population census data of UK. (figure 6-7)

According to what is been observed in the literature, majority of the urban indicators are scaling superlinearly. Indeed, the growth of the cities is measured with the increase in the number of firms. However, with our data we will not agree with this empirical theory. Referring the scatter plot in figure 6 we analyse that the data points are randomly distributed on the slope. With help of scaling law, the exponent is measured as β<1 which is having sub-linear relationship with the firms and population. However, we shall argue against the literature that with increasing of number of manufacturing sectors in England and Wales the growth in size of the city is proportional. This industry example could be considered as limitation in urban scaling theory which shall be explained in the further chapters ahead.

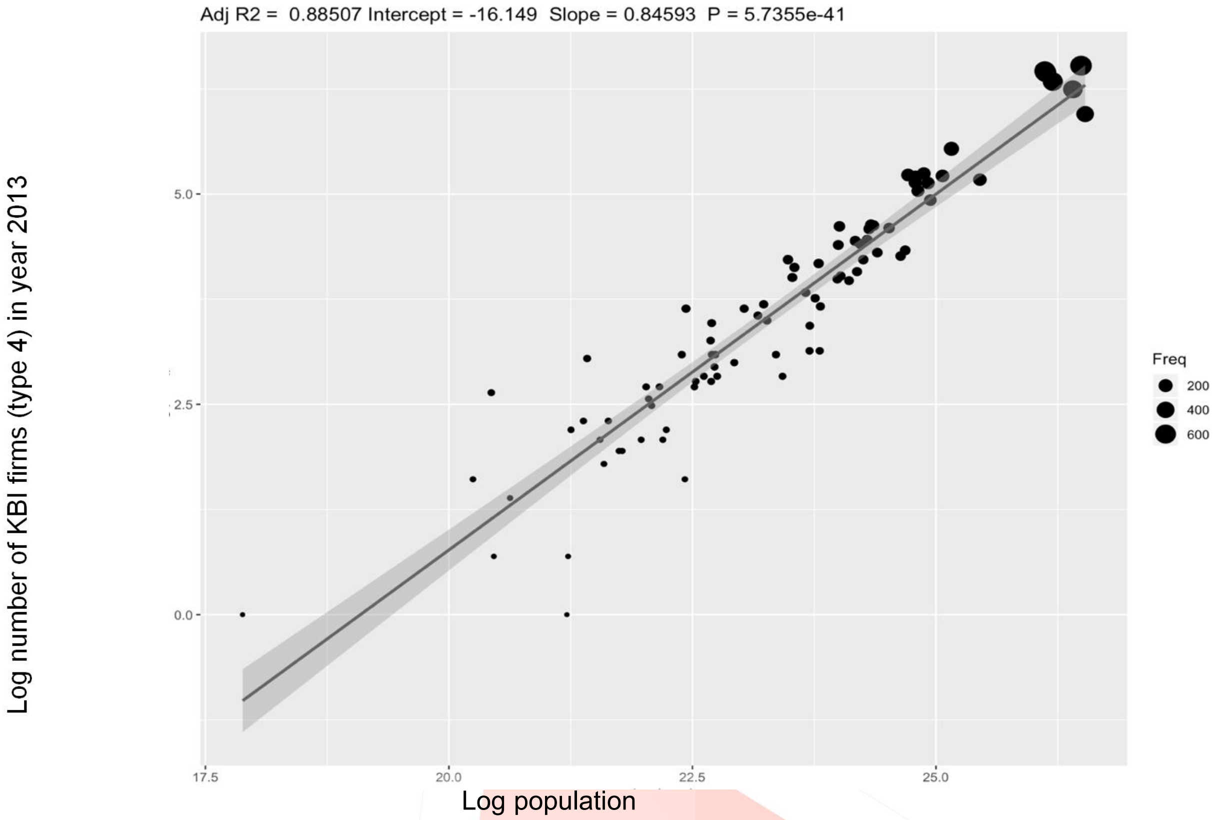


Figure 6.Scatter plots for Knowledge based industry: Other manufacturing (type 4) company house data for the UK.

```
Call:
lm(formula = log(total_m$Freq) ~ log(total_m$pop), data = total_m)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.79350 -0.26573  0.05331  0.26651  1.50246 

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)      -16.14946    0.77679  -20.79   <2e-16 ***
log(total_m$pop)   0.84593    0.03323   25.45   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4848 on 83 degrees of freedom
Multiple R-squared:  0.8864,	Adjusted R-squared:  0.8851 
F-statistic: 647.9 on 1 and 83 DF,  p-value: < 2.2e-16

> coefficients(model_m)
     (Intercept) log(total_m$pop) 
     -16.1494558        0.8459319 
> confint(model_m, level=0.95) # CIs for model parameters
                       2.5 %      97.5 %
(Intercept)      -17.6944530 -14.6044586
log(total_m$pop)   0.7798303   0.9120334
> fitted(model_m) # predicted values
    19201     19209     20110     20120     20130     20140     20150     20160     20170     20200     20301 
 3.988800  3.046829  2.273351  3.053505  4.489143  3.902938  2.438504  4.171159  1.299975  3.900946  4.591465 
    20302     20411     20412     20420     20510     20520     20530     20590     20600     25210     25300 
 3.075287  4.731574  3.041741  4.367554  1.793502  2.900582  2.818279  5.055554  2.503520  3.609439  1.804054 
    25400     26513     26514     26520     27110     27120     27200     27310     27320     27330     27400 
```

Figure 7.Log-log OLS model results for number of industry and city sizes of the UK.

The reason for using this empirical theory of scaling laws in England and Wales data was to create a framework to predict the knowledge-based industries of cities with a set of scaling relations that grow as per its city sizes. Bettencourt-West theory observed ‘As the cities grow bigger, their real income will also increase’ analysed for US metropolitan areas (MSA) evaluating urban planning strategies and scaling exponents (Bettencourt, 2013).

What is knowledge-based industries in the UK and why is it considered?

In this study, we need to find the slope from company house data using the empirical theory of urban scaling laws. So, when we have a detailed look on the data of the UK, we locate industries in the urban region which tend to generate higher incomes and also increase rate of job opportunity.

Here we focus on high tech economics which can play a major role in agglomeration economic of scale. That is, we are interested to analyse the variation within the high-tech economics between cities and its surrounding areas. While analysing the Knowledge based industry from the company house data in the UK, we have firms with five main classifications digital technologies, life science. Healthcare, publishing, manufacturing and other scientific firms. We consider that how these industries relating to economic issues value add in the size of the cities in the UK. This data will help us to identify similar economics showing different patterns of inequalities in cities like Oxford, Cambridge, London or Manchester. We should note that, although these locations are identified as high-tech zones, their characteristic may be insufficient with their economics performance in the rest of the UK.

Explanation on why the knowledge based and leisure industry is explored have a variety of perspectives. We study the particular features which have concentration of information in clusters which shows growth, explain the market demand in the UK and providing a link between them. We review the main features of the literature and understand the scaling behaviour. This includes the spread of KBI, and Leisure in UK and the measurement estimated from β describing the problem and answering the solution.

As per the previous research from Bettencourt – West and E.Arcaute – M.Batty on the scaling exponents theory and their limitation. This paper will examine how knowledge-based industry change over time from 2013 to 2018 and will highlights problems associated with the time in the scaling laws. We shall argue on the key understanding and concept of city size which will be linked with the agglomeration economics of scale.

This paper first discusses about the growing clusters of number of KBI against the city sizes and later on the time frames of the analysis. The quantitative empirical theory will provide findings from the data which will be compared with each other. Finally, summarise our conclusions on scaling linearly or non-linearly.

What is the type of slope we expect from the data? Explain the scaling behaviours analysis:

- ***a. Knowledge-based industry and Leisure industries observation (number of Scientific services firms) reflected as per the city sizes***
- ***b. The difference and changes with regard to time 2013 and 2018.***

To explain this question, we examine the log-log regression model as in figure 8. Earlier we defined the power law scaling equation $y = (mx^{\beta})$ wherein β the scaling exponent is the slope; y is the number of firms (urban indicator) and x is the population size of the city, and m is the constant/intercept. The slope may be positive, negative or neutral totally dependent on data. Indeed, the three types of slope defined as $\beta > 1$ is Super linear, $\beta = 1$ is Linear and $\beta < 1$ is sublinear relationship. In particular, to address the questions, the answers are summarised in table 3: For log-log model, we calculate the scaling exponent β, the interval (β - b, β + b) where b is the error bar in 95% confidence interval of β creating a fluctuation in the data.

SCALING RELATIONSHIPS FOR STANDARD INDUSTRIES IN 2013 FOR THE U.K. VERSUS POPULATION SIZE

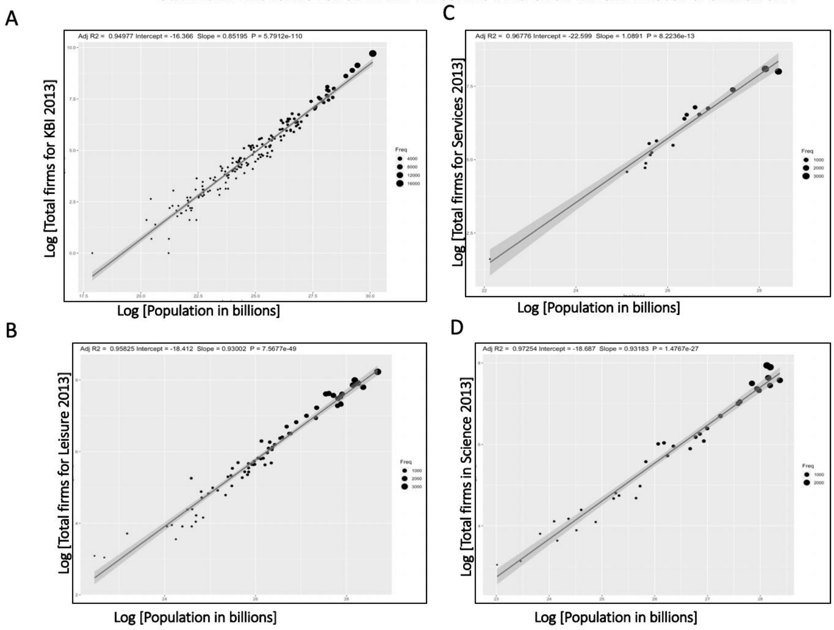


Figure 8.Scaling relationship for standard industries in 2013 vs Population Size in the U.K. (A) Total firms for KBI 2013 (B) Total Firms for leisure Industry 2013 (C) Total firms for Services 2013 (D) Total firms in Science 2013. Best- fit relations of scaling

Table 5.Urban Scaling relations with estimated values of England and Wales cities defined using both knowledge-based industry and Leisure industries.

Y	β	95% CI	Adj-R^2	Observations
Year 2018				
KBI Industry	0.85	[0.82,0.88]	0.94	168
Digital	0.87	[0.84,0.90]	0.97	18
Science	0.91	[0.85,0.94]	0.98	13
Publishing	0.93	[0.87,0.98]	0.97	35
Other Manufacturing	0.84	[0.77,0.91]	0.88	85
Scientific activities services	1.08	[0.98,1.19]	0.96	17
Leisure	0.93	[0.88,0.97]	0.95	71
Food	0.85	[0.71,0.98]	0.94	12
Retail	0.95	[0.89, 1.01]	0.95	43
Entertainment	0.94	[0.84, 1.04]	0.96	15
Year 2013				
KBI	0.80	[0.76,0.83]	0.93	168
Digital	0.89	[0.82,0.95]	0.97	18
Science	0.76	[0.65,0.87]	0.95	13
Publishing	0.93	[0.87,0.99]	0.97	35
Other manufacturing	0.74	[0.68,0.81]	0.86	85
Scientific activities services	1.03	[0.93,1.13]	0.96	17
Leisure	0.93	[0.88,0.97]	0.95	71
Food	0.83	[0.70,0.95]	0.95	12
Retail	0.96	[0.88,1.02]	0.96	43
Entertainment	0.98	[0.88,1.07]	0.97	15

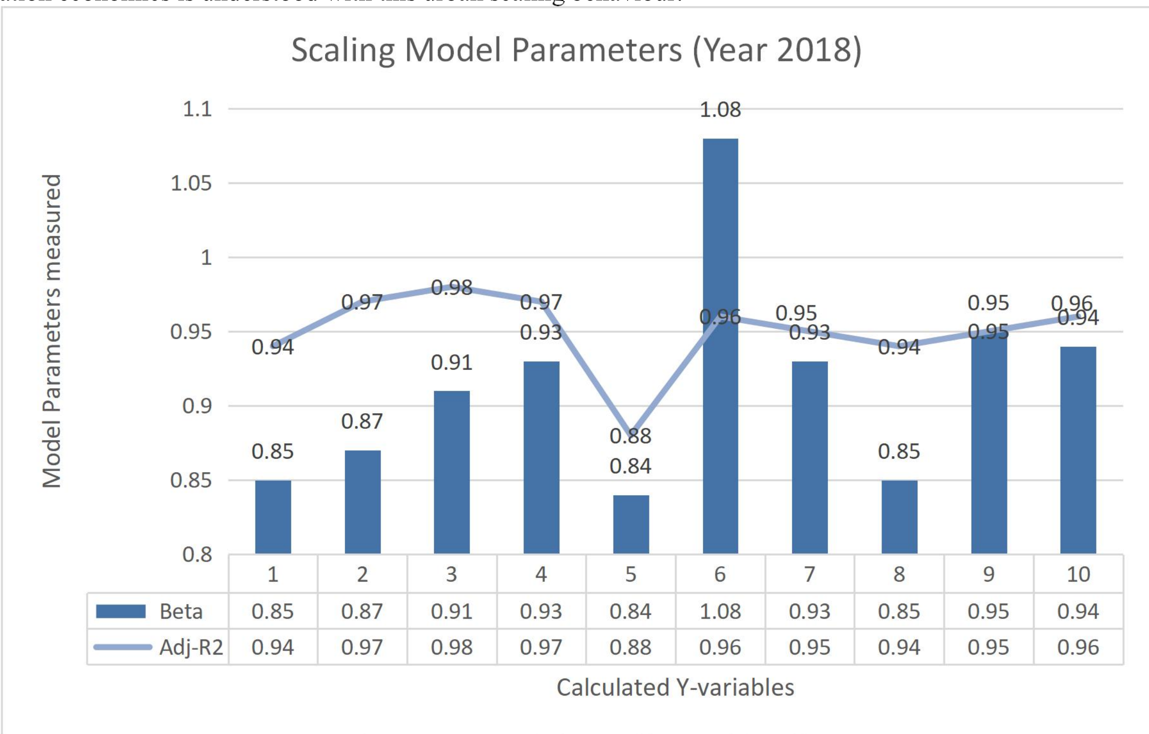
The entire table represents the measure of scaling exponent β both for the knowledge-based industry and the Leisure industry. The β value obtained from the log- scale regression model is reported in the first column. The CI 95% is computed along with adjusted r-squared values higher than 85% with their total observations. The table is self-explanatory which was created using R programming language. Code is pushed in the GitHub mentioned in the appendix.

The result clearly says that the slope is sublinear for knowledge-based industry whereas Other services in KBI, such as technology, innovation or creativity scale super-linearly about 1.08 which increased the average wealth. Indeed, Data science and data digital are close proximity for 1.01 as 95% confidence intervals touches the super-linearly from the linear relationship of scale. This behaviour pattern intensely evaluates that there is a social dynamic that underlines the type of companies, linking them in an integrated network of urban system in a city.

We observe that from the scaling measures, β being linear is extremely mild which is not validating with the scaling laws. For example, the maximum findings for the KBI and Leisure have been the sub-linear slope, by the digital technology, by the life science and healthcare, by the Publishing and broadcasting and by other manufacturing industry. Indeed, the super-linear slope is positively resulted in scientific industry for KBI sector. In the next section, we will discuss about scientific industry, because it has provided a clear example of the main issues that arise in the growing cities to derive generic rules for urban indicators in England and Wales.

Scientific services industries

The result of the other scientific activities of the scaling exponents is super-linear. The number of manufacturing industries from air transport - transport of passengers or freight by air or via space, Architectural and engineering activities; technical testing and analysis, Research and experimental development on social sciences and humanities, natural sciences and engineering, Specialised design activities Post-secondary non-tertiary education, Tertiary education and Post-graduate level higher education is increasing growth with the size of the city. Furthermore, with the results with both knowledge based, and leisure industries and occupation contribute to be observed overall sub-linearity. Not just the size of the city but the agglomeration economics is understood with this urban scaling behaviour.



	1	2	3	4	5	6	7	8	9	10
Beta	0.85	0.87	0.91	0.93	0.84	1.08	0.93	0.85	0.95	0.94
Adj-R2	0.94	0.97	0.98	0.97	0.88	0.96	0.95	0.94	0.95	0.96

Figure 9 Residuals for urban scaling regression model output which is represented in Bar and line graph.

The result is interpreted and well summarised in the line graph shown by figure 9, with the calculated y-variables versus the β ,exponent and R2 model summary data.

In order to examine the importance of scaling exponents β, we study with two parts. The first part is considered with the largest cities of England and Wales plotted against the scientific industries of the analysis. In this case, we have scientific industry type 5 with β = 1.08 which lies in super linearity category. So, we create a scatter plot to show the cities with big size like London does not have the highest value compared to Liverpool and Birmingham. That is, the productivity in Liverpool for industries like freight transport, architectural activities, testing and analysis is highly rooted if compared with cities like London and Leeds. (refer figure 10.)

In the second part, we notice there is lack of resemblances with the scaling exponents when compared with the big cities of different knowledge-based industries for year 2013 and year 2018. So, it is normal to see some unexplained to scale in some places. Different empirical theory for location may cause different results. However, from the results using the scaling analysis, the Slope in most knowledge/leisure-based industry is sub-linearly scaled which is closely related to linear scaling distribution. Scientific activities industry is often seen to scale super-linearly having similar characteristics of urban indicators, while cities with bigger size locations are not very similar in their behaviour of scaling relationships.

Similarly, in order to verify the relationship between the scaling parameters for year 2013- 2018 we display the results on the bar chart and scatter plots.Additionally, we observe that most of the measurement in the scaling relations, there is a mild fluctuation in the exponent variables form 0.88-0.97 which falls under the sublinear and very few in super-linear slope bucket. Referring the result table 4, we should also note that the difference and changes with regard to time 2013 and 2018 is minor within the five years scaling behaviour for both the knowledge based and leisure industry. Indeed, Science industry scaling exponent β have increased from 0.74 to 0.96 still belonging to the sub-linear scaling relationship. Hence, the difference between the time frames is not much that affects the slope in the data.

With these two above sub questions, we answer the research question for the scaling behaviour of the knowledge-based industry. In the next section, we will discuss the scaling theory, explore the limitation and create an outline for the future research.

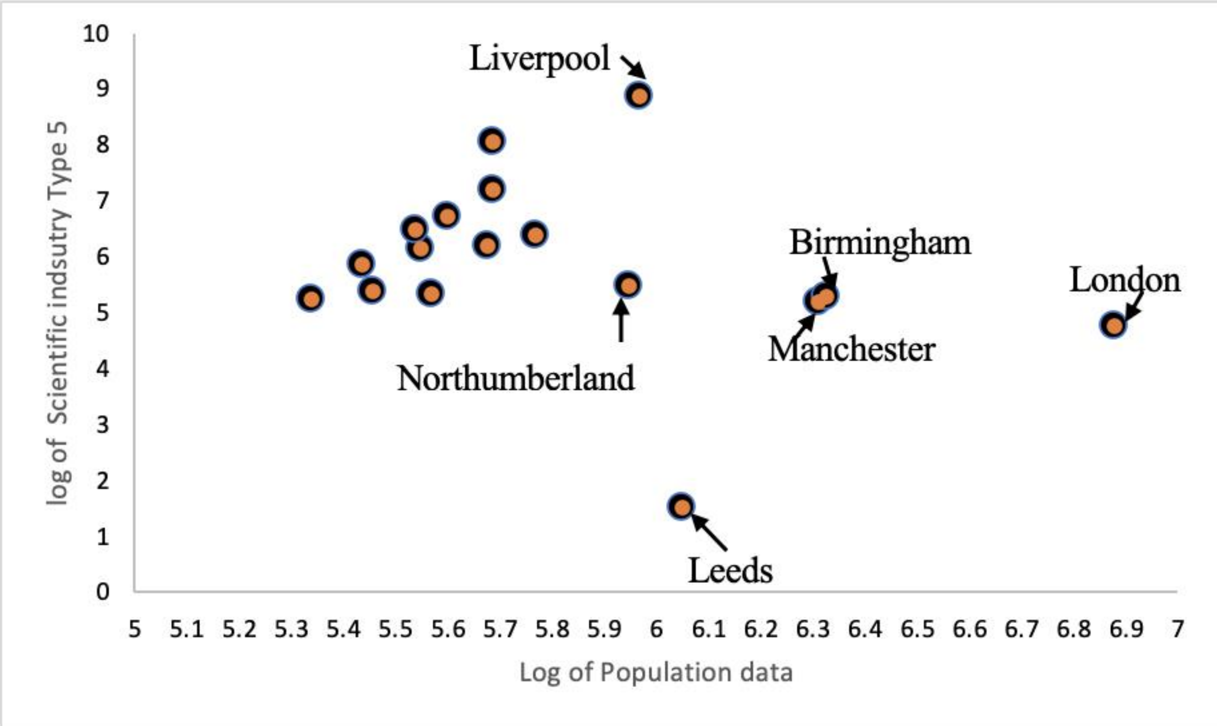
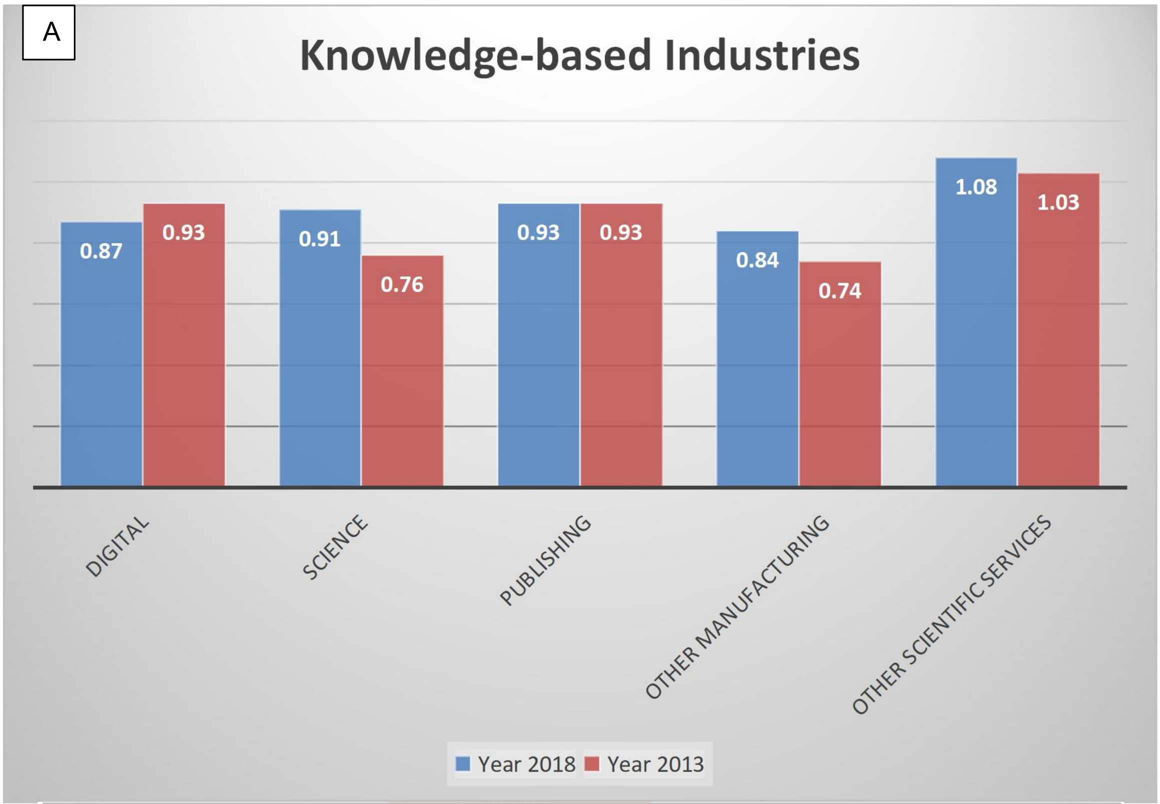
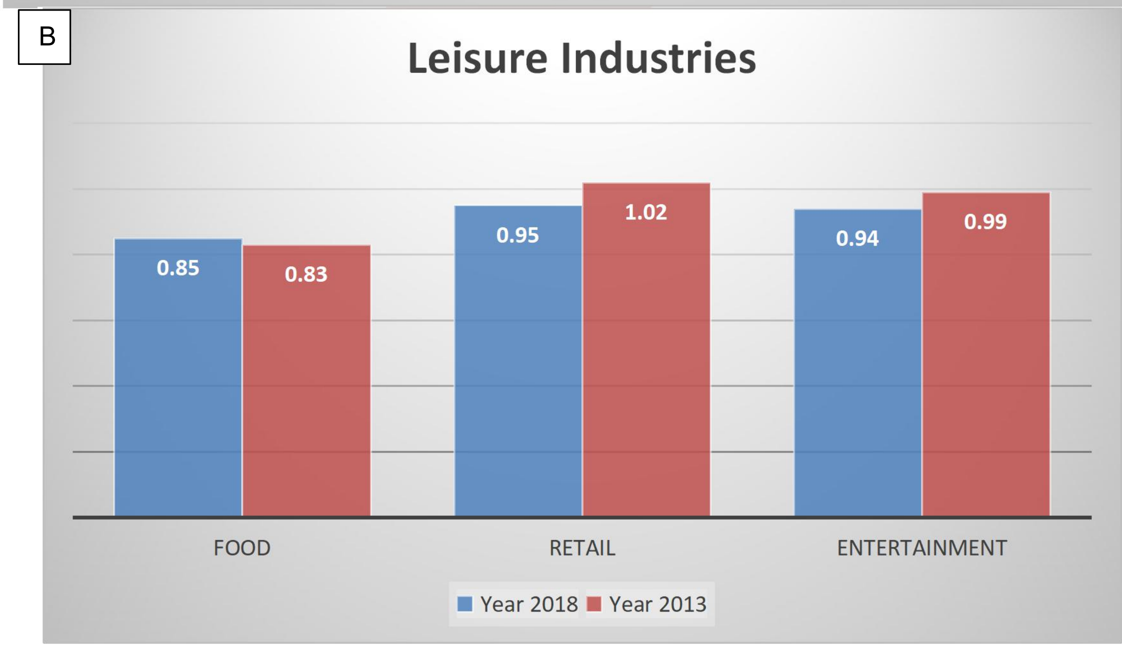


Figure 10.Distribution of cites for superlinear scaling relationship which is identified for scientific industries type 5 in company house data for the UK.

Testing changes of Industries classification

To understand the differences with the scaling exponents result obtained, the taxonomy between the two categories of industries are compared. With the measures transformed variables in figure 11.A to 11.D, the bar, pie and line plot represent the variation and changes in both the industries. This dictates that the $\beta>1$ with the superlinear slope, that is increasing growth in the city size for industries such as scientific services of KBI and retail of Leisure industry (Figure 11A,11B): $\beta=1$ linear meaning number of firms is related (equal.) to population of the city. Lastly $\beta<1$ with the sub-linear slope, outputs such as digital, science, publishing, food, entertainment is resulting slow growth with size of the city in England and Wales of the UK. (Figure 11C,11D)


A
Knowledge-based Industries
0.87
0.93
0.91
0.76
0.93
0.93
0.84
0.74
1.08
1.03
DIGITAL
SCIENCE
PUBLISHING
OTHER MANUFACTURING
OTHER SCIENTIFIC SERVICES
Year 2018
Year 2013


B
Leisure Industries
0.85
0.83
0.95
1.02
0.94
0.99
FOOD
RETAIL
ENTERTAINMENT
Year 2018
Year 2013

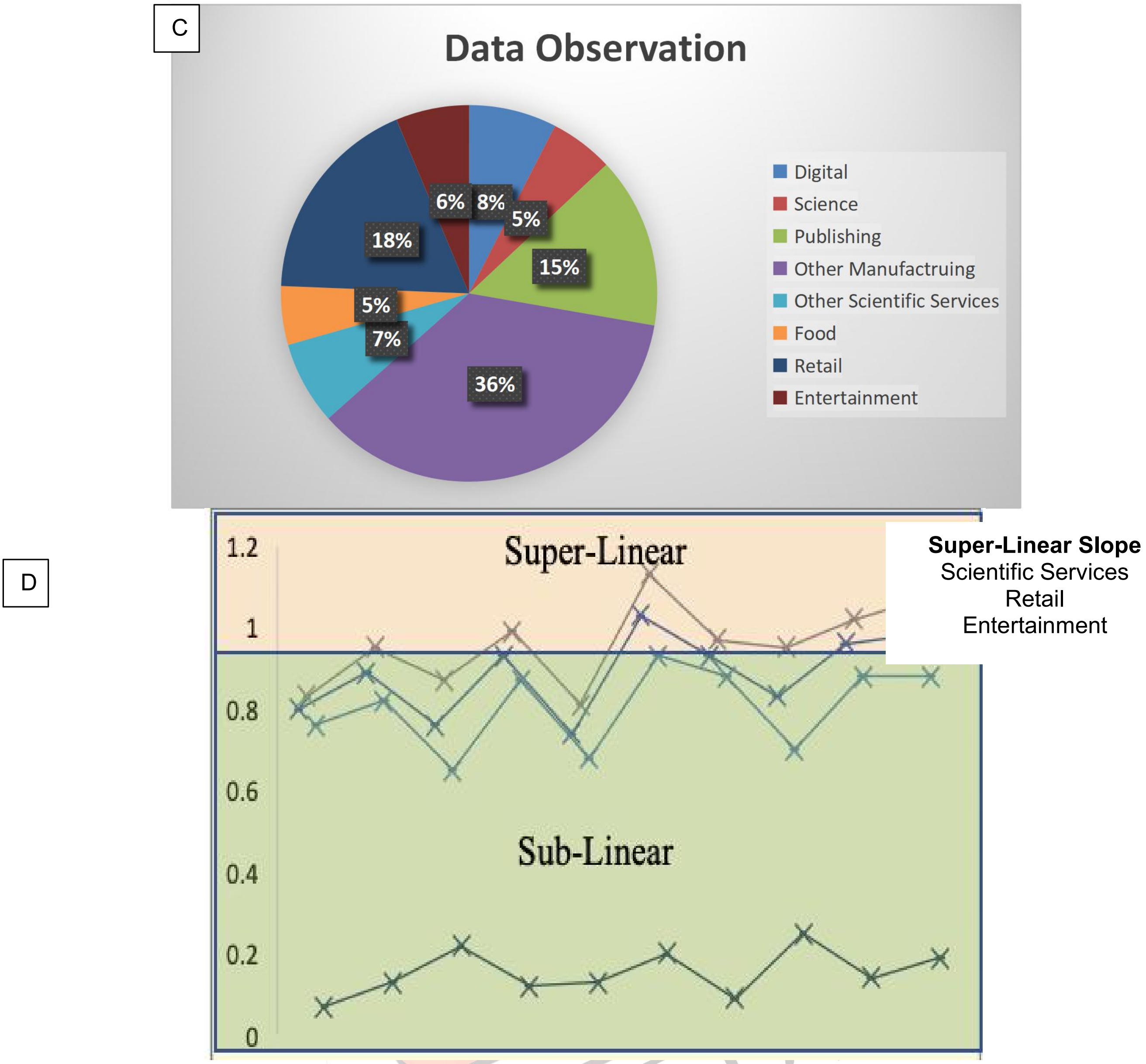


Figure 11 Fluctuation of scaling exponent generated from log-log model output. (A) Knowledge based Industries for year 2013~2018 (B) Leisure Industries in the UK for year 2013 ~ 2018 (C) Company House data with 239 data points for the UK industrial classification

V. DISCUSSION

Research findings

In the above analysis, we explore the urban scaling Laws in knowledge-based industries and Leisure industries using the company house data. We also explore the relationship between them. From the perspective from our literature, we define agglomeration economics with regard to their city sizes as explained and implemented by many policy makers. Therefore, when we analyse the overall rule of cities sizes and urban indicators using the urban scaling empirical theory, we must separate the data of the two. In addition, due to the particularity of their scaling behaviours, our focus is towards the knowledge-based industry and leisure industry concerned with population size for England and Wales. The analysis found that the slope in the data has sub-linearly measured except in scientific industry which has super-linearly categorised. That means, there is an increase rate of education, technical testing and analysis, in places where the growth of cities is faster compared to those who are slower, by population size. This is a general rule of the urban built environment: that city is stronger with its economic hub compared to its integrated urban system. Bettencourt- West theory has not expanded their research in England and wales, but this study is evident and much inclined to E. Arcaute - M. Batty paper on deconstructing the scaling laws. In addition, our scaling analysis with the large set of industrial company house data, each exponent has been changing with mild fluctuation pattern. Through the three research questions, it is found that England and Wales cities does not rule the empirical theory and with our findings Liverpool in the UK is evident to play a highly productive city for these knowledges based scientific industry which have scaled super-linearly. Although there was a discrepancy that as per the size of the city we will see the growth in their urban indicators; London does not follow this within the urban system of the UK. The results obtained are highly independent with our data and are not very integrated to any system or theory. It is difficult to judge whether this method will provide a linear relationship with much better results. What is more, the geography of cities is continuous, and even though we can generally know the land types and main human activities in each region, it is difficult to specify a boundary. So, it is normal to see some unexplained scaling relationship in some places. Different definition of cities with urban agglomeration behaviours may cause different results. However, from the results of the measuring scaling laws behavior, the knowledge-based industries types of the classifications are closely sub-linearly related to their city distribution as per population size. Indeed, Scientific industries were super-linear – varied linearly with the expected size which changes in different locations with scaling characteristics.

Why do we expect the exponent in knowledge-based industry is super-linear, some of them as sub-linear and some industries as linear? However, we shall not agree with the literature and there is an enormous misconception on this theory especially for England and Wales data.

Research significance

As data analyst, we need to consider and solve the problem of how we can predict the size of the cities with various urban indicators used in agglomeration economics of scale and maintain the urban system operational. The main problem faced by the urban system theory is the imbalance between cities. This imbalance is reflected in the two dimensions of time and space; unbalance growth of cities in every time period and unbalanced growth scaling at different locations of the world. Therefore, the understanding of this empirical theory of urban scaling laws can better help to predict statistical measurement and understanding their difficulties which will help to tackle the rigorous approach in defining the cities boundaries with scaling in cities. Although the literature helps us to create a framework and influence the cities growth, our research can provide a general understanding of all industries especially knowledge-based growth in UK cities.

In addition, we adopt scaling relations to predict the characteristic of a city realising the industries' as urban indictors scale with their city sizes. In (Bettencourt et al., 2007) and (Cottineau et al., 2015), they carried out scaling law analysis summarising variations in the sizes of the city with the United States MSA cities and France respectively, which also revealed the spatial aggregation of local units measuring the order of magnitude of scaling estimations. However, what they used was population census data, infrastructural, socio-economic data collected from local gov level, so the variables used by their empirical theory are the number of properties in the city applied to the urban system. The principle is to compare the similarity of the set of scaling relation between the cities in proportion to the number of properties available. However, there is a problem with this approach: if there are two urban indicators (properties), one with the income (GDP growth) of the city and the other is urban infrastructure (for example the area of the road) but these same order of magnitude will be used in other cities in the same period of time, this will result in one city to scale super-linearly and the other city will have sub-linear relationship. That is to say, the scaling laws available for the two cities with similar scaling behaviours will be very different, leading to scaling exponent clustering into different classes. In our study, we use the company house data, which can calculate the number of industries both KBI and Leisure for each city in England and Wales more intuitively. This solves the problem of population census data for cities, because the similarity of their observed prediction is more convincing than their defined scaling relations, and the slope results could be more accurate.

Research limitation

We must admit that our research also has some limitations:

Firstly, the results obtained from the measurement of the scaling exponents using the log-log scaling method is not a statistically robust method and provides uneven numbers which are different to predict the movement of cities definition. As per the literature study, we don't find an accuracy measure for the exponent from the data. For a more rigorous analysis, we should refer to a variety of data for cities with different location. In this way, not only the samples are more convincing, but also the scaling relationship between the urban metrics and cities sizes can be observed.

Secondly, we introduce the hierarchical properties of cities at different levels of agglomeration economics so called urban scaling laws. These methods require us to determine the number of urban industries like the knowledge-based industry, and we cannot automatically select the specific number. In our research, we mainly determine the numbers of firms through practical experience and observation of the characteristics of each type of firms. However, this method is not rigorous enough. If more firms are selected, another result of scaling exponent and slope may be obtained. Therefore, more methods should be compared in the defining the cities with its order of magnitude.

Thirdly, in terms of variables, we used the population census data along with the company house data with specific industries over year 2013-2018. Although this is the main factor that affects the measure and scatter plot of the results, their scaling behaviour of each city is also influenced by many other factors, such as the population density of area, location of the KBI firms, transportation and connectivity, infrastructure, economic hub of the city etc. and this convenience of other approaches used for constructing a city. However, these factors were not taken into account in this study.

Finally, the aggregated levels of scaling laws as explored in literature are having difficulty in this study and leading to sub-linear results.

VI. CONCLUSION AND FUTURE WORK

With our urban system becoming more complex, data on the cities will become ubiquitous. This empirical theory is only a platform to create new techniques that will be necessary to collate and analyse large amount of data produced in the real world. This study took place in England and Wales as a case we study that Liverpool is scaling super-linearly with the scientific knowledge-based industries and to study the statistical pattern of the city boundaries. The results show that industries of the urban system are closely linearly related to the size of the cities. This will be of great help to the future work for defining the city boundaries and exploring the relationship as a follow-up research.

While some of the results in this paper fit our intuitions about the patterns of the cities in England and Wales, and other details are also very instructive. For example, in US MSA data, the income grows more than the city size and resulted in super-linear scaling.(Bettencourt et al., 2007) When we analyse this theory for our cities, we cannot provide standard results compared to Bettencourt. When we try to define using our industry indicators against the city, we do not get super-linear scaling. The bigger the city in the size does not really proof that it has more income, more jobs or its richer in England and Wales. London is a special city so in out study it observed as Outlier.(Arcaute et al., 2019) We believe that similar analysis can be extended and combined with other urban infrastructures and different cities around the world focusing on India or China cities.

In the future, we will try to use a variety of models to predict the city size with its order of magnitudes as defined by urban scaling laws. Indeed, we could work on this empirical theory and analyse to create a template, because the prediction in this study in-between KBI and Leisure industry and size of cities lack the practical significance. We will combine the trend to analyse the historical data to form the predicted variables. We will analyse the impact of other variables like the demographic data, time series, climatic, sustainability data for the study of the size of the city. A comparative study of this analysis will provide important insights in the growth, innovation and scaling of the cities. We will use nonlinear scaling relations to estimate and predict the exponents, indeed, to differentiate in their city's economics and other factors of growth.

To optimise the city scaling relationship, we can also plan to use the predicted exponents results to rebalance the distribution of the urban system in the built environment. However, the scatter plot and the detail study does not convey any much evidence on the agglomeration effects in cities of England and Wales.

Within the scope of quantitative analysis, there is another direction that we have not studies in depth: How the data for human activities will change with the addition, deletion or relocation of the size of cities. We intend to continue this work in the future through a comprehensive comparison of scaling power law and their results in the observed scaling behaviours.

Appendix: Code availability

The data have been collected from the open source website and have no data privacy concerns. This report has no additional data. We have built and analysed the mathematical models. This project has no competing interest. Data is uploaded on the GitHub in the paper repository using command line. Link here. https://github.com/architectjyothi?tab=repositories

REFERENCES

[1] Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., Batty, M., 2014. Constructing cities, deconstructing scaling laws. J. R. Soc. Interface 12, 20140745–20140745. https://doi.org/10.1098/rsif.2014.0745

[2] Arcaute, E., Hatna, M, 2019. Scaling laws: insights and limitations 21.

[3] Batty, M., 2008. The Size, Scale, and Shape of Cities. Science 319, 769–771. https://doi.org/10.1126/science.1151419

[4] Bettencourt, L., West, G., 2010. A unified theory of urban living. Nature 467, 912–913. https://doi.org/10.1038/467912a

[5] Bettencourt, L.M.A., 2013. The Origins of Scaling in Cities. Science 340, 1438–1441. https://doi.org/10.1126/science.1235823

[6] Bettencourt, L.M.A., Lobo, J., Helbing, D., Kuhnert, C., West, G.B., 2007. Growth, innovation, scaling, and the pace of life in cities. Proc. Natl. Acad. Sci. 104, 7301–7306. https://doi.org/10.1073/pnas.0610172104

[7] Cottineau, C., Finance, O., Hatna, E., Arcaute, E., Batty, M., 2018. Defining urban agglomerations to detect agglomeration economies. Environ. Plan. B Urban Anal. City Sci. 239980831875514. https://doi.org/10.1177/2399808318755146

[8] Cottineau et al., 2015. Diverse cities or the systematic paradox of Urban Scaling Laws | Elsevier Enhanced Reader [WWW Document]. https://doi.org/10.1016/j.compenvurbsys.2016.04.006

[9] Gill, I.S., Goh, C.-C., 2010. Scale Economies and Cities. World Bank Res. Obs. 25, 235–262. https://doi.org/10.1093/wbro/lkp022

[10] Glaeser, E.L. (Ed.), 2010. Agglomeration economics, National Bureau of Economic Research conference report. The University of Chicago Press, Chicago.

[11] Leitão, J.C., Miotto, J.M., Gerlach, M., Altmann, E.G., 2016. Is this scaling nonlinear? R. Soc. Open Sci. 3, 150649. https://doi.org/10.1098/rsos.150649

[12] Pumain, D., Paulus, F., Vacchiani-Marcuzzo, C., Lobo, J., 2006. An evolutionary theory for interpreting urban scaling laws. Cybergeo. https://doi.org/10.4000/cybergeo.2519

[13] Sarkar, S., 2018. Urban scaling and the geographic concentration of inequalities by city size. Environ. Plan. B Urban Anal. City Sci. 239980831876607. https://doi.org/10.1177/2399808318766070

[14] West, G.B., 1997. A General Model for the Origin of Allometric Scaling Laws in Biology. Science 276, 122–126. https://doi.org/10.1126/science.276.5309.122

[15] West, G.B., Woodruff, W.H., Brown, J.H., 2002. Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. Proc. Natl. Acad. Sci. 99, 2473–2478. https://doi.org/10.1073/pnas.012579799