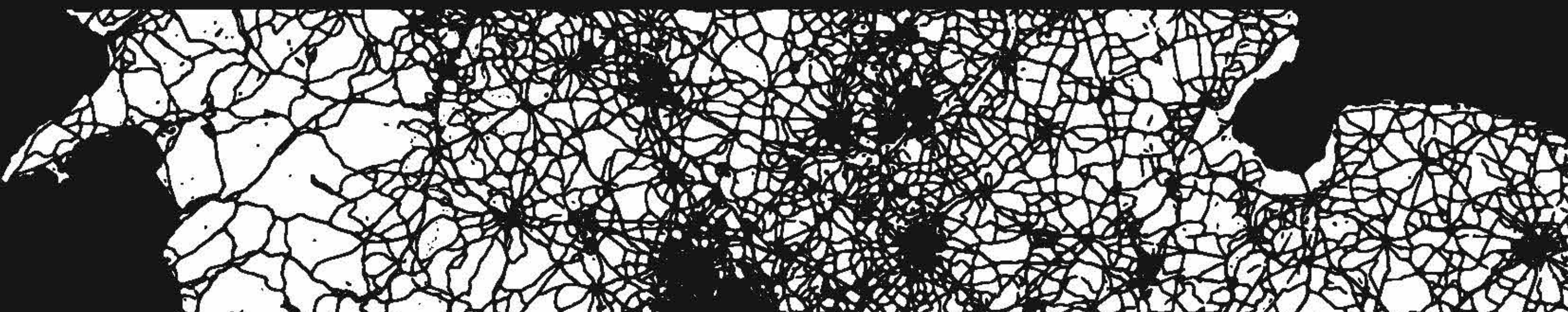
Urban Productivity & Spatial Patterns Across Scales

HADI ARBABI

SHEFFIELD - FEBRUARY 2019



A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

The University Of Sheffield.

Department Of Civil & Structural Engineering.

RISE | Resources, Infrastructure Systems, and built-Environments

Urban Productivity & Spatial Patterns Across Scales
A Multi-Scale Exploration of Urban Networks and Their Hierarchical Configurations

RISE – Resources, Infrastructure Systems, and built-Environments Research Group
Department of Civil & Structural Engineering
The University of Sheffield

Submitted February 2019, Defended June 2019, Revised September 2019

Sir Frederick Mappin Building
Mappin Lane, Sheffield
S1 3JD – UK
Phone 0114 222 5728
h.arbabi@sheffield.ac.uk
www.envbldg.sheff.ac.uk

Most artisanally typeset in LaTeX

Note for Americans and other aliens:
Milton Keynes is a new city approximately halfway between London and Birmingham.
It was built to be modern, efficient, healthy, and, all in all, a pleasant place to live.
Many Britons find this amusing.

— Neil Gaiman and Terry Pratchett

To my parents and MB…

It is so certified here that all contained within this volume and presented hereinafter are the author's own work except where it is clearly referenced to that of others.

H. E. A.

Statement of Conjoint Work

The body of work comprising the bulk of this volume has resulted in the following peer-reviewed publications:

1. Chapter 4 – Arbabi, H., Mayfield, M., Dabinett, G. (2019). Urban Performance at Different Boundaries in England and Wales through the Settlement Scaling Theory. *Reg Stud*, 53(6), 887-899.
2. Chapter 5 – Arbabi, H., Mayfield, M., McCann, P. (In Press). Productivity, Infrastructure, and Urban Density—an Allometric Comparison of Three European City-Regions across Scales. *J R Stat Soc A Stat*.
3. Chapter 6 – Arbabi, H., Mayfield, M., McCann, P. (In Press). On Development Logic of City-Regions: Inter- Versus Intra-City Mobility in England and Wales. *Spat Econ Anal*.
4. Appendix A – Arbabi, H., Mayfield, M. (2016). Urban and Rural—Population and Energy Consumption Dynamics in Local Authorities within England and Wales. *Buildings*, 6(3), 34.

The author confirms that his contribution to these jointly co-authored papers with the supervisory team – Martin Mayfield (Civil & Structural Engineering), Gordon Dabinett (Urban Studies & Planning), and Philip McCann (School of Management) – covers the entirety of the research design, analysis, and primary interpretation of the results and overall amounts to at least 80% of the total work in each paper.

Where a chapter includes direct reproductions of previously published material, these are those of the author's sole contribution and are indicated at the outset of the chapter.

Abstract

Understanding the nuances at play across different spatial scales is of crucial importance when considering urban economic-energetic size-cost performance, specifically when longer-term consequences are considered. Through the application of an allometric understanding of cities, a more nuanced narrative is offered highlighting the interplay of urban productivity and spatial configurations of human interactions across scales. This is presented in three parts.

In the initial examination of the urban economic-energetic size-cost balance across spatial scales, we seek new insights on the effects of scale in relation to urban connectivity and density for maximizing urban size-cost balance. For this, we use the urban system in England and Wales as a topical testbed where agglomeration-based arguments have been used in support of better inter-city connectivity in order to address a historic North-South regional economic productivity divide. The inadequate connectivity thought to be affecting the economic performance across the urban network in England and Wales, however, is shown to permeate across spatial scales. More broadly, this points at a scale-induced hierarchy of urban connectivity concerning potential improvements needed at inter- and intra-city scales.

This is followed by an examination of the universality and transferability of scaling insights, and their nuances, between different cities and systems of cities. Considering the current transport schemes designed to address the North-South economic gap, we examine the continental comparisons drawn specifically from the inter-city transport infrastructure connecting the Randstad in the Netherlands and Rhine-Ruhr metropolitan region in Germany. Our examination points towards fundamental differences that exist in the structure and distribution of population density across the countries and their city-regions across various scales. Additionally, the cross comparison demonstrates that, although scaling insights are transferable between urban systems, a simple multi-scale assessment of individual systems of cities in isolation is sufficient when investigating urban connectivity from an urban allometric point of view.

Finally, returning full circle to the effects of spatial scales and distance on the geographical patterns of urban connectivity, we review a mathematically grounded approach to sort and organize the intra- and inter-city connectivity hierarchy while matching complementary infrastructural needs based on size-cost balances for a number of different scenarios.

Together, this narrative provides a somewhat enhanced and most crucially spatially multi-scale examination of the arguments regarding connectivity and agglomeration in an urban context.

keywords: agglomeration; urban scaling; urban characteristics; infrastructure planning; densification; transport; energy

Figures in this volume are sized and placed according to the amount of time and attention they are to be paid. Margin figures, in particular, are provided for broad patterns rather than details.

Contents

List of Figures

List of Tables

1

Introduction

Material prepared for this chapter have been used in the following–

Arbabi, H., Mayfield, M. (2016). Urban and Rural—Population and Energy Consumption Dynamics in Local Authorities within England and Wales. *Buildings*, 6(3), 34.

In 2009, the world underwent a significant but subtle change. By mid-year 2009, the overall number of people inhabiting urban and metropolitan areas across the world was reported by the United Nations (2014) to exceed that of those still populating rural areas. Over 3.3 billion people at the time, by 2030, the urban population is expected to grow to almost five billion people. In terms of the urbanized spatial extent required to accommodate this rise in urban population, an additional of up to 800,000 square kilometers of the planetary surface is estimated to become covered by the built-up area of cities in the same time period (Seto, Güneralp, and Hutyra, 2012). While this pattern of urban growth appears inevitable, it can be readily attributed to the potentially more favorable environments in urban settings and what they offer in terms of education, employment opportunities, and most other services (Worldwatch Institute, 2007). In the meantime, according to the International Energy Agency (2013), the same human settlements have come to be responsible for nearly 76% of the energy consumption globally. This corresponds to a striking 60% of global fossil fuel consumption and, subsequently, upwards of 71% of the direct energy-related greenhouse gas emissions (World Energy Council, 2013). Hence, the importance and significance of the impact of cities in the context of climate change, their optimality in transforming their energetic consumption to wealth, and the necessary measures to address these have slowly become more widely acknowledged and studied.

The questions and debates surrounding city size and performance, and thus the issue of the desired size of cities, on the other hand, have been a long-standing one in almost all disciplines that concern themselves with the study of cities. This has led to a large body of literature exploring different aspects and issues regarding the city size and form, from:

- studies measuring metabolic activity as it relates to energy and resource consumption and efficiency for cities of different sizes as a function of city structure, shape, and form,
- those conceptualizing cities as complex thermodynamic systems, and
- those exploring appropriate size of cities and its effects on performance through the size and balance of their economies and resources, to
- studies investigating the existence of rank-size distribution rules in cities, and
- allometric[1] scaling of different city properties and characteristics with their size.

[1] Allometry as a tradition in biology is concerned with the study of the link between the growth of body size and shape (West, 2018). In contrast with *isometric* transformations where processes are governed by a constancy of proportions and one-to-one scaling of metrics, *allometric* processes follow a variable rate of growth resulting in a change of proportions.

In their endeavor to define, quantify, and understand cities and the underlying dynamics that drive them, these studies provide insights that can be used to systematically investigate, identify, and debate planning and energetic characteristics[2], needs, and requirements of cities against their theoretically idealized portrayals or comparative amongst themselves. When generalizing cities with respect to their size, the often accepted, yet not unchallenged, wisdom of this body of literature includes observations along the lines of:

[2] For the purposes of this volume, the energetic characteristics referenced are those that pertain to urban connectivity patterns and processes unless specifically stated otherwise.

- bigger cities are more productive economically,
- bigger cities consume more energy and resources, and
- denser cities are more efficient materially and energetically.

Although a less reductive reading of the literature should acknowledge the voices of dissent and the occasional empirical evidence at odds with the above items, *bigger does more with less* when ignoring particular subtleties and nuances. Some of these subtleties, however, could grow in significance when we shift perspectives from obtaining knowledge, expanding understanding, and increasing our explanatory power as it relates to constructing a brand new Science of Cities (Batty, 2013) to their application in real-life diagnostics, examination, and generation of planning and infrastructural solutions. Given the importance of cities socially, economically, and environmentally, examination of such issues could be seen as vital to how cities are planned specifically, and hence to the wider society in general.

1.1. Research scope and questions

There is no shortage of evidence that cities and urban cores are the global nuclei of innovation, wealth generation, resource consumption, and energy dissipation. Against a backdrop of expanding urbanized areas and increasing urban populations with limited resources available to sustain them, the need to design and maintain urban fabric and infrastructures in a manner that enables cities of higher productivity for minimal dissipated resources is hence self-evident. An implicit implication of observations that bigger does more with less in practical terms concerns efforts that would enable a collective of *smaller* urban areas to virtually act as a single *bigger* and hence more productive and efficient poly-centric city-region. Intuitively, one of key levers for this is often identified as the connectivity required between and across such collection of units (Hall and Pain, 2012, pt. 4). In practice, however, the ability to clearly identify, design, and implement connectivity measures, in particular those concerned with transport and mobility, that in fact improve collective performance, for a given definition of performance objective, is dependent on the availability of appropriate models and understanding of the system at appropriate scales of intervention.

One, among many, of the subtleties that impact the generality of connectivity-led agglomeration arguments and the bigger-does-more-with-less principle is the effects of spatial scales on interpretation and application of empirical observations and infrastructural interventions when tweaking and tinkering with cities and systems of cities for better energetic and economic performance.[3] This work, therefore, seeks to explore the effects of spatial scales on the observations of size-induced agglomeration elasticities. For this, we adopt an agglomeration-compatible framework that provides a direct and non-abstract interpretation of urban size-cost performance balance focusing on economic output, urbanized area, energy consumption, and their interrelation through population size at various spatial scales. Consequently, investigating this overarching area, we endeavor to more specifically examine the following broad questions with reference to the potential effects of spatial scales:

[3] While we provide a formal description of 'agglomeration economies' in chapter 2, unless stated otherwise, we use agglomeration and allometric/scaling effects interchangeably in referring to the propensity of denser and more populace urban units to exhibit larger productivities and efficiencies.

A. do allometric arguments regarding connectivity and agglomeration remain valid across different scales and definitions of city boundary?
B. to what extent are such scaling insights transferable between different cities and systems of cities? and finally,
C. what are the effects of spatial scales and distance on the geographical patterns of connectivity-based agglomeration within such scaling frameworks?

To enable an exploration of the aforementioned questions, we use the urban systems in England and Wales, Germany, and the Netherlands as topical case studies.

1.2. Structure and outline

We begin chapter 2 with a broad overview of the variety of approaches and disciplines studying and analyzing cities and urban environments. This broad overview is followed by a more specific examination of the works pertaining to the effects of size, returns to scale, and agglomeration. Over the course of chapter 2, we see an overview of the evolution of the methodologies and arguments employed across disciplines concerned with urban spatial patterns and performance, from urban economics to industrial ecology, culminating and converging in a generalized allometric formulation of cities and systems of cities. We conclude the chapter with positioning of the research questions outlined above in reference with the wider literature examined.

In chapter 3 we review the main allometric framework and the particular model to be used exploring the research questions set out in this chapter along with a description of the data and procedures followed. Following this, chapters 4, 5, and 6 provide the main bulk of the narrative each with their own self-contained description of methodology addressing each of the research questions A to C.

Chapter 4 begins by outlining the relevance of the urban network in England and Wales as a case study. This is discussed in the context of the countries' current infrastructure planning policy as related to the connectivity-led agglomeration arguments. We then proceed by first examining the existence and coherency of population scaling patterns at different spatial scales and categorical delineation of cities for economic output and urbanized area. In doing so, the chapter provides an assessment of urban size-productivity for each boundary definition analyzing density-connectivity optimality within the UK urban network and by proxy identifying infrastructural needs, e.g. increased connectivity or built density. What we will see is a systemic lack of adequate connectivity in a large portion of city units considered at various spatial scales. More interestingly, however, the broader interpretations of the results point at a scale-induced hierarchy of urban connectivity concerning potential improvements needed at inter- and intra-city scales.

In investigating the transferability of these multi-scale insights, chapter 5 extends the analysis to explicitly include a multi-scale comparison of the urban network in England and Wales with its Dutch and German counterparts. We first contextualize this comparison given the emphasis put on continental inspirations, especially those of the Randstad in the Netherlands, and Rhine-Ruhr metropolitan region in Germany, in advocating the connectivity-based agglomeration strategies reviewed in chapter 4. This comparison culminates in the observation of lower densities in the English North as underpinning the size-productivity differences of the city-regions in the three countries. This points towards fundamental differences that exist in the structure and distribution of population density between the countries and their city-regions across various scales. Additionally, our examination also demonstrates that, although scaling insights are transferable, a simple multi-scale assessment of individual systems of cities in isolation is sufficient when investigating urban connectivity from an agglomeration point of view.

Having reviewed the effects of spatial scales on urban performance balance in an isolated urban system and also in comparison with other urban systems, in chapter 6, we close the narrative by exploring the geographical patterns of connectivity-based agglomeration that can emerge within such scaling frameworks using a urban size-cost performance balance as a criteria. On the surface, the chapter provides visually and spatially explicit comparators to the city-regions of our specific case study. More fundamentally, the findings reaffirm observations from previous chapters that there exists a persistent scale-induced hierarchy of urban connectivity concerning potential improvements needed at inter- and intra-city scales which are predominantly frequent and potentially more adequately addressable over short or intra-urban distances.

Lastly, we conclude with an overall discussion of the analyses in chapters 4 to 6, including potential policy implications of the model in the English and Welsh case study, and a brief concluding chapter summarizing the findings and outlining available avenues for future work in chapters 7 and 8.

2

Science of Cities

Material prepared for this chapter have been used in the following–

Arbabi, H., Tan, L. M., & Mayfield, M. (2018). Comments on ‘A multi-level framework for metabolism in urban energy systems from an ecological perspective’ by Pulido Barrera et al. (2018) *Resour Conserv Recy*, 136, 463-465.

We begin our review of the literature with a bibliometric overview of works concerning cities. This allows for a mapping of the 'current' disciplinary approaches, the extent of their interdisciplinary engagement, and ultimately aids with identifying interdisciplinary knowledge gaps and/or emerging approaches that could be exploited as a way forward. To enable this systematic review of the preceding literature fundamental to this work and before fleshing out those directly instrumental to the methodology, we first employ a co-occurrence analysis to examine the broad intellectual space dedicated to the study of cities. The rest of this chapter first provides a brief description of the bibliometric technique followed by an overview of the disciplines studying cities, their methodologies, and overall insights they provide. We then further delve into those particularly concerned with urban size and returns to scale before concluding with a brief summary of the subsequent research needs addressed in this volume.

2.1. A bibliometric overview of the literature on cities

A cursory search for the keywords *city*, *cities*, and *urban* on the Web of Science™ (WoS) offers over half a million records flagged by the database. Table 2.1 shows the top six research areas and the record count corresponding to each area. Demonstrably, cities as a subject matter and/or embedding context provide a focal point for several distinct disciplines from Social Sciences to Engineering. Given the specific domain of interest laid out for this work, the search can be further refined with specific modifiers: *productivity* OR *mobility* OR *density* OR *agglomeration* OR *transport* OR *infrastructure*. This substantially reduces the number of unique records to around 34,000.

Table 2.1: Top research areas and their corresponding record count for a search on *city* OR *cities* OR *urban*.

Research Area	Record Count
Environmental Sciences and Ecology	102,661
Engineering	84,066
Public Environmental Occupational Health	51,545
Urban Studies	40,992
Business and Economics	36,446
Geography	31,191

Note that the six areas presented constitute over half the records corresponding to the keywords searched.

A co-occurrence analysis highlights the distance and relation between different records based on the frequency of their co-occurrence in the cited references of other publications, see Figure 2.1. As such, two publications are considered connected if they appear in the references of another record with the weight and strength of their connection assumed proportional to the number of records in the references of which they are observed (White and Griffith, 1981). Along with direct and weighted-direct citation analyses, co-citation analysis can help discover research themes and interdisciplinary dialogs between them (Persson, 2010). Using the Bibexcel shareware (Persson, 2014; Persson, Danell, and Wiborg Schneider, 2009), an initial investigation was made into the coupling of the main research areas. Figure 2.2 shows the the co-occurrence of the broad disciplinary areas where the size of each node is proportional to the number of records tagged under that area and the edge width proportional to the frequency of each pair appearing together. It is clear that the study of cities, even when constrained to a limited number of keywords, is of an inter- and cross-disciplinary nature although the strength of these collaborations appears to differ from pair to pair.

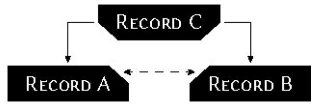


Figure 2.1: Schematic showing a co-citation link between records A and B where each are cited by record C.

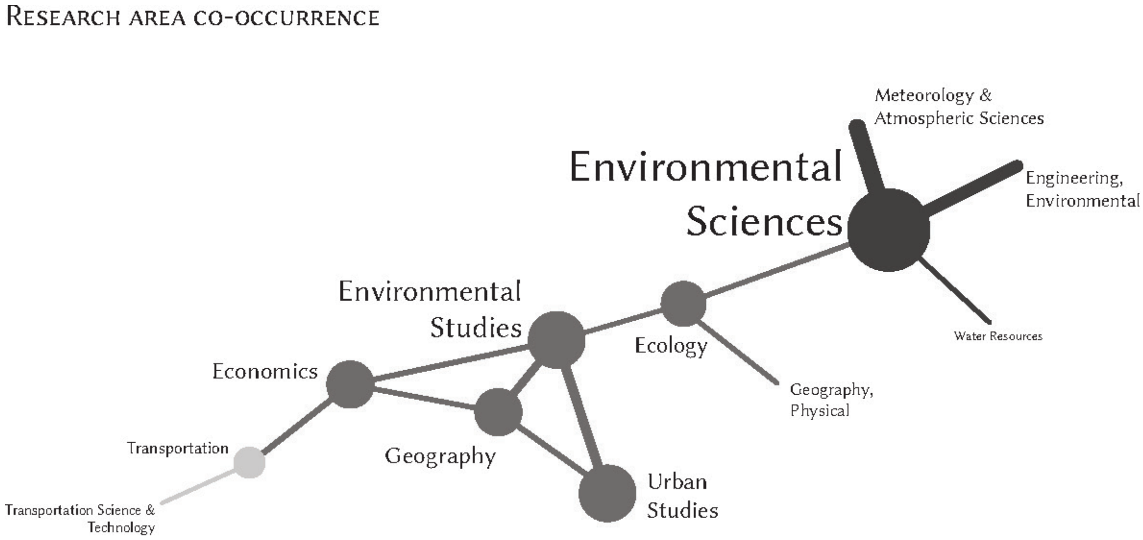


Figure 2.2: Network representation of the top ~5% research area co-occurrence among the WoS records – Note that edge-width is proportional to co-occurrence frequency with f_{min} = 156 and f_{max} = 692.

The co-citation network of the records helps develop an overarching picture of the literature exploring cities and the various methods and approaches that are currently being used. It also facilitates the visualization of the connection and dialog between these differing approaches and research areas. Figure 2.3 shows the co-citation map developed. A brief examination of the clusters reveals a similar structure to that observed in Figure 2.2.[1] The three most populated clusters at the bottom involve papers broadly studying the energetic and material consumption within cities through three distinct quantitative methodologies and conceptualizations of consumption in ecological and engineering themes while the smaller clusters to the

[1] Note that the network has been trimmed for legibility. As such fewer social sciences and urban economics records are visible compared with the engineering records that are inherently more frequent in publication and citation.

Co-citation map of urban research

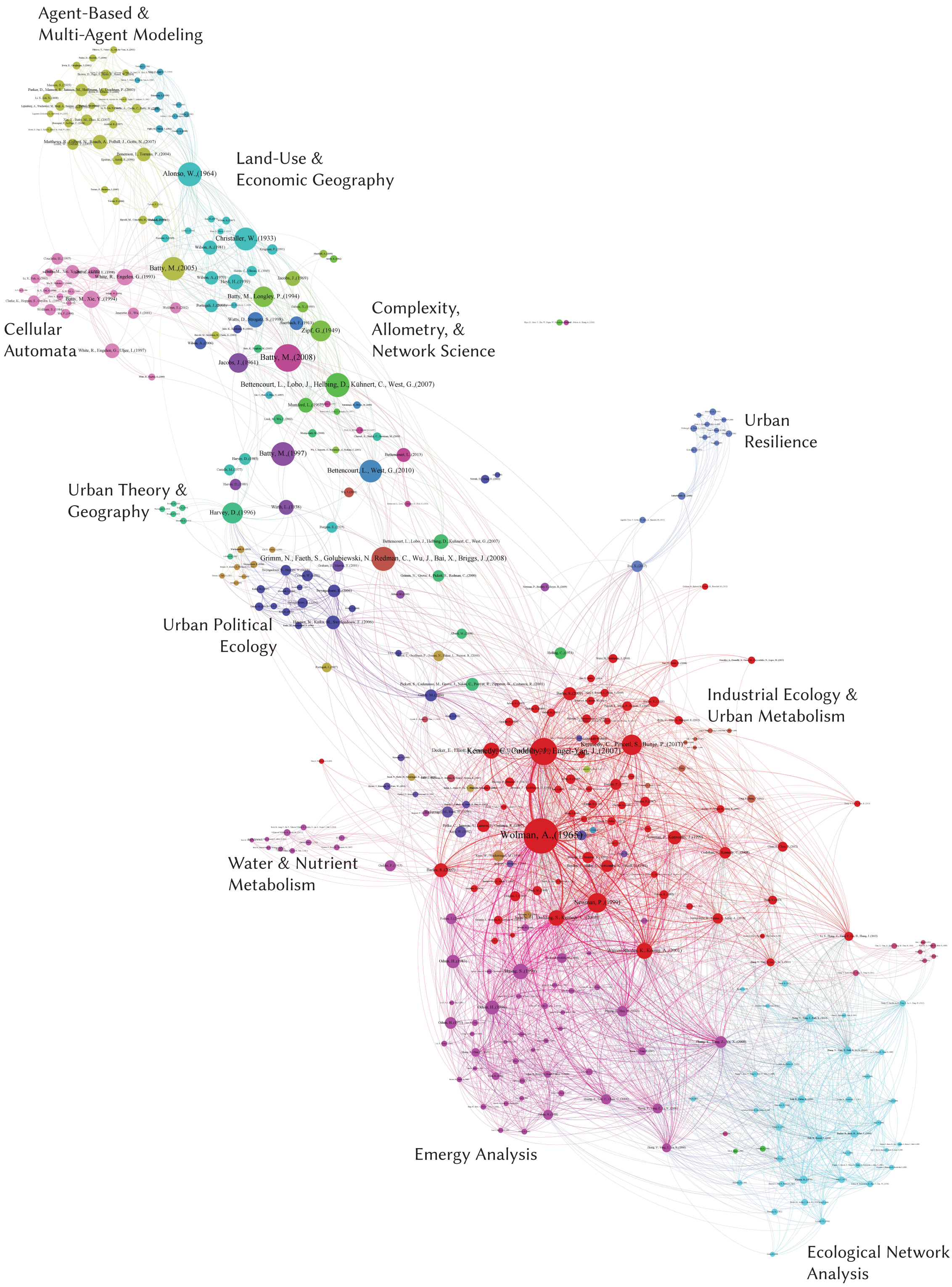


Figure 2.3: Network representation of the co-occurrence among the records – Note that the map only shows co-occurrences with $f \geq 5$. The bottom half is dominated by the engineering and ecology disciplines while the top clusters belong to a wider variety of disciplines. Clusters correspond to specific research themes and/or methodologies edges.

top are comprised of papers exploring different aspects of cities and their properties from the perspective of more socially orientated disciplines, e.g. sociology, political ecology, geography, economy, etc., and using different methodological tools such as multi-agent modeling, cellular automaton, rank-size and allometric analysis, etc. Despite the variety of methods and disciplines dedicated to the study of cities, the network also illustrates the disconnect between some of these thematic clusters and also the individual works and/or authors who have managed to provide a bridging link between such clusters. The following will briefly discuss the major themes from the prominent clusters providing an overview of the overall methodological approaches and frameworks.

2.1.1. Cellular automata and land-use and transport models

The top corner of Figure 2.3 is populated with publications concerning shape and geometry of cities and their patterns of development as it relates to various mobility and economic principles. Thematically, the larger cluster consists of two densely connected communities dedicated to the application of modeling methods, i.e. techniques largely based on the use of cellular automata and agent-based models, to the study of transportation access, land-use patterns, and urban growth. The theoretical and intellectual core of these studies, however, traces back to a slightly older community which includes the main urban spatial economic theories and models and provides the intellectual tie connecting the other two method-based communities. These range from those concerned with settlements size and patterns of development, that is following Christaller's *Central Place Theory* (1966) which sought to formulate the spatial distribution and hierarchy of human settlements on a flat ever-expanding domain based on their market, transport, and administration optimality, and those exploring land-use patterns within cities, e.g. Burgess' *Concentric Zones Model* based on descriptive observations of land-use patterns around the central business district in Chicago (1925) during the 1920s, complemented more mathematically by Alonso's *Bid-Rent Theory* (1964) which linked the likely land-use in each zone to the land-value attributable to each use based on economic activity and distance from the central business district, and variations of the main concentric model to adjust for socio-geographical features and transport routes (Hoyt, 1964; Hoyt and United States Federal Housing Administration, 1939) and to accommodate patterns with more than one specialized central district (Harris and Ullman, 1945), see Figure 2.4 for an abstract representation of these models.

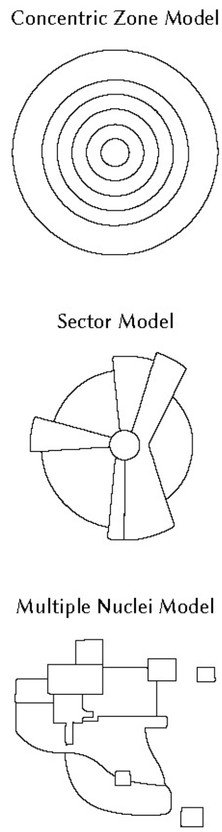


Figure 2.4: Schematic representation showing the concentric, sector, and multi-nuclei models.

Cellular automata are a class of discrete models which in a two di-

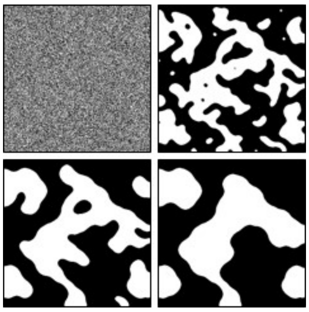


Figure 2.5: Schematic of a simple implementation of cellular automata and its evolution for a 432×432 board where the color of each cell is decided based on the prominent color in its neighborhood – Clockwise from top left, random seed, $t = 0, 50, 300, 650$ to system at stability.

mensional space are comprised of a regular array of cells with the state of every cell at each time interval decided based on a transition function factoring in the state of the cell and its neighbors (for various definitions of neighboring cells) at the previous time interval(s). They have been shown capable of mimicking complex behaviors seen in physical systems (Chopard and Droz, 1998, ch. 2; Hoekstra, Kroc, and Sloot, 2010) where simple rules and interactions at fine resolutions lead to emergent behavior across the overall analytical space (Wolfram, 1984). This may be most demonstrably seen in the simplest implementation of cellular automata in studying segregation patterns where a binary transition rule determining the value of each cell according to the value of a majority of its neighbors leads the system from a random seed to organized and ordered patterns, Figure 2.5. As it pertains to cities, cellular automata has been a re-occurring model in studies that aim to investigate temporal dynamics of change and growth in the urban fabric with the notable majority modeling and simulating land-use type and cover changes using different transition decision frameworks, from simple implementation of those put forward in the above-mentioned theoretical models and heuristics regarding land-use development to more complex transition rules incorporating machine-learning techniques to drive rules based on time-series empirical land-use data (Barreira-González, Gómez-Delgado, and Aguilera-Benavente, 2015; Elmenreich and Fehérvári, 2011; Feng, Liu, and Batty, 2015). It should, however, be noted that despite the powerful ability of cellular automata to mimic complex generative systems and its application in exploring and uncovering underlying land-use dynamics, for instance the existence of temporally resilient spatial relationships between different land-uses despite socioeconomic changes over time, the quality of its implementation is dependent on the transition criteria used which still lack vigorous theoretical development (Stanilov and Batty, 2011).

Similar to cellular automata, agent-based modelss also work by studying system evolution based on the interactions between smaller individual components the behavior of which is modeled accordingly. These components, i.e. the *agents*, could be representations of individual urban inhabitants or even larger entities and stakeholders, e.g. utility companies, financial firms, agents with different modes of transport, etc., depending on the scope and objectives of the study and the overall model. Unlike the cells in CA however, the agents of agent-based modelss are not necessarily geographically bound and are thus capable of integrating stochastic locational dynamics more explicitly. This has resulted in integration of cellular automata and agent-based modelss in modeling land-use and urban growth patterns where the interactions of agents from the agent-based models framework, be it individuals or representative institutions, e.g. developers, businesses, farming co-ops, etc., influence the transition rules of the cellular automata representing properties of the physical land parcels often coupled with geographic information system (GIS) data (Fisher-Gewirtzman and Blumenfeld-Liberthal, 2012; Xie, Batty, and Zhao, 2007). In a broader context of cities and their energy consumption and infrastructure, agent-based models frameworks have also been adopted for exploring urban electricity, energy networks, and smart grid implementation using goal-based decision criteria for the agents representing various stakeholders and consumers (Gonzalez de Durana, Barambones, Kremers, and Varga, 2014; Rylatt et al., 2013).

2.1.2. Industrial ecology and urban metabolism

Complementing the studies mostly dealing with spatial patterns of urban economy and its geography, the opposite end of the network in Figure 2.3 is comprised of studies that are concerned with the consumption and flows of resources in cities. Wolman's seminal study (1965) in which he estimates consumption and circulation values for fuel, water, waste, and pollutants for a hypothetical American city of a million inhabitants, is widely acknowledged to have ignited the interest in evaluating the flows of materials and energy streams that move in and out of cities and hence the metabolism of cities, now conventionally described as '…the sum total of the technical and socio-economic processes that occur in cities, resulting in growth, production of energy, and elimination of waste…' (Kennedy, Cuddihy, and Engel-Yan, 2007). Since Wolman, the field of industrial ecology has seen a string of studies and projects focusing on the evaluation of all or subsets of these flows for a number of cities, e.g. Hong Kong (Boyden, 1981; Newcombe, Kalma, and Aston, 1978; Warren-Rhodes and Koenig, 2001), Sydney (Newman, 1999; Newman et al., 1996; Newton et al., 2001), Taipei (Huang, 1998; Huang and Chen, 2009; Huang and Chen, 1990; Huang and Hsu, 2003), Lisbon (Deilmann, 2009; Niza, Rosado, and Ferrão, 2009), etc. While the accounting methodologies and frameworks used in these studies have changed and evolved throughout the years, from simple estimations based on national averages à la Wolman to more sophisticated, disaggregated, and at times dynamic models (Bergsdal, Brattebø, and Müller, 2014; Kazanci and Ma, 2012; Rosado, Niza, and Ferrão, 2014; Tanikawa and Hashimoto, 2009), they share a common use of metaphors and analogies likening the processes of an urban system to the metabolism of an ecosystem or an organism and a philosophy rooted in systems thinking. These studies can be broadly grouped based on the core approach they adopt in evaluating urban metabolism of different cities, i.e.

- those mainly engaging accounting practices, e.g. material flows accounting/analysis (MFA), to benchmark and measure city flows and overall efficiencies for a conceptualized, often black-box, systems representation of the city (mainly corresponding to the central red cluster and its peripheries, Figure 2.3),
- the group incorporating an emergy[2] approach to model the city and the ecosystem in which it is embedded (purple cluster), and
- the network-based studies which use methods developed by the wider environmental ecology community, e.g. network environ/ecological analysis (NEA), to investigate the interrelation between urban subsystems when consuming and producing different flows (light blue cluster).

[2] Emergy or *embodied energy* is a concept used in ecology which seeks to measure and harmonize the totality of both direct and indirect energy streams embedded in a given service or product with the energy quantity harmonized in terms of the equivalent planetary solar input required.

The majority of the urban metabolism studies utilize the MFA framework in one way or the other. The methodology relies on a steady-state understanding of the streams of energy and materials throughout the urban components. Under these assumptions, one can attempt to measure and evaluate flows between components applying mass-balance to the in- and out-flows and stocks, be it material or energetic, for each component. In these studies, the output of the accounting practice is then often used to address waste generation and optimize material circulation by closing the loops within the urban subsystems. A common obstacle in this area, however, appears to have been a lack of uniformity in the materials and flows considered (Kennedy, Stewart, Ibrahim, Facchini, and Mele, 2014) along with inconsistent boundaries used for the city and its subsystems (Ramaswami and Chavez, 2013; Ramaswami, Chavez, and Chertow, 2012). This is compounded with the sparsity of recorded data on the desired flows which inevitably results in high levels of uncertainty (Patrício, Kalmykova, Rosado, and Lisovskaja, 2015).

The emergy concept, as conceived by Odum (1974, 1996), although not exclusively devised to study cities, applies the same systems thinking to ecosystems and the cities within them while trying to account for the thermodynamic differences in the quality of the energy entering and leaving the subsystem processes. In doing so, the emergy literature, and by extension its less environmentally/ecologically oriented counterpart exergy literature (An, An, Wang, Gao, and Lv, 2015; G. Chen and Qi, 2007; Koroneos, Nanaki, and Xydis, 2011), aspires to quantify every flow, be it energetic, material, or monetary, in terms of its solar energy equivalent, i.e. embodied energy, hence emergy, which can be thought of as the primary form of the planetary energy input.[3] Compared with the multi-unit analyses of the MFA, in theory, the emergy approach addresses and rectifies the problem of incomparable flows between different subsystems and improves consistency by accounting for the losses and efficiencies of different transformations taking place within the system. Realistically, however, devising meaningful or accurate unit conversion factors for many of the transformations that take place within urban areas and the ecosystem surrounding them poses methodological challenges and remains impractical (Y. Zhang, Yang, and Yu, 2015).

[3] Although related, exergy is a more formally defined thermodynamic property that describes the maximum *useful work* that can be extracted from a given process or product given the environmental conditions of its surroundings.

Finally, network analysis provides a methodological basis to further analyze system models that are assembled using generic MFA or emergy approaches. The method uses the formalization of the Leontief's Input-Output analysis (Fath and Patten, 1999) where the interdependency of the subsystems and flows are investigated using matrix representation of their network and through the application of graph theory. The use of the network analysis enables an extended exploration of the direct and indirect effects of different subsystems and their synergistic relations beyond the simple accounting exercises of the MFA studies (**tan_ecological_2018**; Fath and Borrett, 2006; Li, Zhang, Yang, Liu, and Zhang, 2012; Yang et al., 2014). Furthermore, this family of approaches can establish trophic hierarchies based on the flow contributions of each node to the rest of the network or vice versa and as such provide a basis for drawing comparisons between sector hierarchies within urban metabolic structures and those of more balanced and self-sustaining *natural* ecosystems gaging self-sustenance in urban systems. More recently, similar network based analyses have been applied in a more spatially explicit contexts studying the transformation of land-use types through time. These examine changes in the trophic consumption, production, and accumulation in and over different land-use patches, e.g. urbanized land, forests, grasslands, etc., in lieu of the traditional flows of the conceptualized sectors, e.g. primary and secondary energy producers, consumers, etc. These also investigate the overall emission savings or losses associated with change from one land-use to the others (Y. Zhang et al., 2016).

Nevertheless, the combined body of literature across these three thematic approaches, especially those revisiting the same city during different time periods, have been fundamental in highlighting the imbalances in material and energy extraction and consumption pointing to a continued increase in the throughput required to drive cities and the waste generation and emissions exported out of cities.

2.1.3. Urban ecology and urban political ecology

One of the main points of criticism raised with and within the previously described body of literature, from now on industrial ecology/urban metabolism, is the omission and disregard of the interrelations at finer spatio-temporal scales in the frameworks used to quantify urban metabolism as studies focus on generalized accounting practices at the expense of the identification of the role played by various social actors and the dynamics of their effects on the ecosystem. As addressed within the studies belonging to the industrial ecology/urban metabolism clusters, these criticisms are framed in the forms of calls for inclusion of the human element and the extension of the indicators studied to those expressive of socio-economic processes, mostly through interdisciplinary and collaborative studies engaging academics and methods from appropriate disciplines (Kennedy, Pincetl, and Bunje, 2011; Newman, 1999; Pincetl, 2012). The same gap in industrial ecology/urban metabolism is, however, regarded very differently from the viewpoint of those further removed from the community expressed either as industrial ecology/urban metabolism's misapplication of the ecosystem- and organism-based metaphors (Golubiewski, 2012a, 2012b) resulting in a reductionist and simplistic picture of modern urbanity and its complexity (Gandy, 2004) or its inability to *theorize* as to the nature and dynamics of socio-environmental processes (Swyngedouw, 2006).

While the criticisms that the industrial ecology/urban metabolism has historically stretched the application of metabolic metaphors by often ignoring the differences in scale and function that exist between organism, ecosystems, and human-made systems in favor of the similarity observable in a general flow of *things* and thus become reductionist and simplistic in its approach is ultimately true, they are also partial and biased to some extent due to disciplinary understandings and practices. Mostly stemming from a life sciences background, this framing of the critique is more critical of the application of the language used which limits the interdisciplinary reach of understanding cities and urban environment as ecological ecosystems. As such, the main studies of cities sympathetic to this viewpoint often tend to be a part of long-term projects with high aspirations to capture a full picture of the complexity and interrelation of not just the socio-economic but also detailed ecological, biophysical, and biochemical components of the urban landscape. The Long Term Ecological Research Network (2018) funded by the National Science Foundation of the United States is perhaps one of the more prominent examples accommodating such projects specifically as a part of their Baltimore Ecosystem Study and Central Arizona – Phoenix

LTER which are embedded in urban areas and include community outreach and integration programs. Although these types of projects can be viewed as ideal in understanding the systematic intricacies and hierarchies of interactions that exist throughout different scales between the multitude of urban components, from the cycles of simple chemical nutrients, such as phosphor and nitrogen, vegetation and patch dynamics to emissions associated with human activity, they require distinctly long periods of active and continuous monitoring and data collection and the involvement of large numbers of practitioners from varying disciplines to over-come the industrial ecology/urban metabolism shortcomings (McPhearson et al., 2016).

An alternative articulation of these criticisms belong to the urban political ecology community the intellectual background of which is more closely tied to the qualitative branches of geography and sociology.[4] Unlike the industrial ecology/urban metabolism community, the focus in UPE is less so the quantity of the *things* moved into, out of, and within cities but rather the institutional processes that cause and direct such flows. As such, their utilization of the metabolic metaphor concerns itself with the processes through which humanity and nature affect one another in the urbanized space (Brenner, Madden, and Wachsmuth, 2011; Heynen, Kaika, and Swyngedouw, 2006; Swyngedouw, 2006) and is influenced by the earlier engagements of such metaphors within the Marxist literature (Newell and Cousins, 2015). The concerns of the community are then the inter-relationships and dichotomies that are manifest in the struggle between the social and political powers in the appropriation and production of nature and urban space (Wachsmuth, 2012). An example can be seen in Gandy's (2004) observation of water infrastructure and metabolism across cities. In a traditional industrial ecology/urban metabolism manner, a study would typically concern itself with the amount of water, energy, and other nutrients that are involved within the metabolic systems model of a city and whether or not relative to the population and through time this metabolism has behaved efficiently whilst Gandy, and by association the UPE intellectual tradition, would make an extension and pay more attention to the fact that the most efficient water utilities appear to also be those under public control and not privatized and are consequently more interested in the power struggles and inequalities inherent in socio-political mechanisms and repercussions of such observations.

[4] This can be visually confirmed by investigating co-occurrence edges present in Figure 2.3.

Not surprisingly, there have been efforts made in the space between the quantitative industrial ecology/urban metabolism and the qualitative and discursive urban political ecology where, despite being fully organized, the accounting-based methodologies of the former are either put in context and linked to the social, historical, political, and geographical characteristics of the study space or are adjusted to communicate the effects of city metabolism on the environment and the planet more explicitly. However, it should be noted that these examples, specifically those concerned with using more environmentally oriented indicators, such as ecological footprint[5] and human approriation of net primary production (HANPP),[6] tend to expand their focus to boundaries larger than single urban environments and cities to regions and sometimes entire countries (Blomqvist et al., 2013; Kastner, Erb, and Haberl, 2015).

[5] Ecological footprint can be defined as the amount, i.e. area, of untouched productive ecosystem, be it land-based or marine, needed to sustain the human resource consumption and absorb the subsequent waste (Wackernagel, 1994).

[6] HANPP provides a measure of the effects of human induced land-use change on the primary production, i.e. the total energy produced by living components of an ecosystem, e.g. through photosynthesis (Haberl et al., 2012).

2.1.4. *Complexity, allometrics and a new science of cities*

Stepping away from the well-defined community clusters explored so far, the connecting bridge between them comes in the form of topically diverse yet thematically linked research. These more or less attempt to form an understanding of this science and physics of cities as permeating not only some but most of a city's characteristics. Alternatively, they attempt at codifying and formulating these characteristics, including but not limited to urban micro-economics or energy/material metabolism, through more sophisticated and sometimes non-equilibrium physical analogies with running themes of complexity, entropy, and self-organization. The observations of the underlying and seemingly persistent rank-size distributions among cities, and components therein, of an urban system could be traced back to the works of those including Auerbach[7] (1913), Figure 2.6. The popularization of a complexity-orientated conceptualization of and approach to cities, however, is more commonly attributed to Jane Jacobs. Jacobs (1961, pp. 442-458), influenced by works of Weaver (1948), mapped the kind of problems cities manifest as those with organized complexity where multiple interconnected and intertwined variables effect the dynamic of the city and are affected by changes in each other through various intricate hierarchies. This, not surprisingly, hearkens back to the systems thinking within industrial ecology/urban metabolism albeit with more of an emphasis on size, shape, and growth – economic, infrastructural, or otherwise – rather than the industrial ecology/urban metabolism's focus on pure metabolic energy and/or materials. Although topically more diverse compared to the other intellectual communities discussed previously, the strain focusing on the nature of size relationships in area and volume of cities and their infrastructure, including the derivation and application of self-similar geometric patterns in understanding and modeling urban growth and development processes (Batty and Longley, 1994; Y. Chen, 2014; Feng and Liu, 2015; Terzi and Kaya, 2008), has been re-occurring.

[7] Such distributions were later popularized by Zipf, and are commonly referred to today as Zipf's Law. Batty (2008) provides a concise perspective on size and scaling in city planning connecting together many of the topics mentioned here.

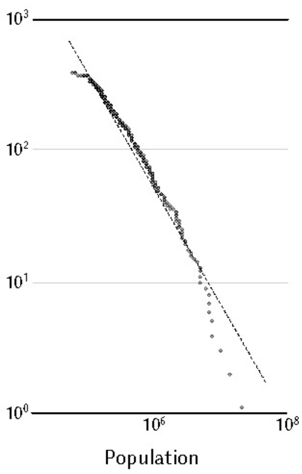


Figure 2.6: Double logarithmic plot showing city rank against population for 383 US metropolitan statistical areas in 2017 – data from United States Census Bureau (2018).

2.2. Size and returns to scale

Following a bigger-does-more-with-less line of reasoning, larger urban areas are thought to be associated with higher resource/infrastructural efficiencies and economic productivities. As hinted at in the previous chapter, these arguments could carry the implicit implication that a collective of *smaller* urban areas made to, say through better inter-unit connectivity, act as a virtually single *bigger* poly-centric city-region can then benefit from these size-related productivities and efficiencies. Empirical observations of such patterns, although not necessarily conclusive, are abundant across the thematic disciplines reviewed in the previous section. With the majority of the disciplines attempting to understand mechanisms that drive these agglomeration benefits, one of the recurring patterns across the mechanisms explaining this higher comparative productivities and efficiencies of larger cities concerns the various influences of urban connectivity and density. From the perspective of the more energy/resource-oriented studies, the scale efficiencies of bigger conurbations are rooted in their density and physical morphology (Mohajeri and Gudmundsson, 2014). Meanwhile, within the more geography/economics-oriented studies, the two are assumed to be instrumental by facilitating the mixing of people, ideas, and goods (Glaeser, 2010).

2.2.1. Energetic arguments and observations

In the late 1980s', transport studies concerned with potential energetic efficiencies of urban areas saw a surge and refocusing of arguments around urban population and built density. Newman and Kenworthy (1989, 1999) studied the energy consumption in transportation systems within cities. Their study investigates the variations of the transport energy use as a function of population density of several major urban zones globally. They note that the annual consumption of fuel for transport follows an inverse power-law with respect to population density, Figure 2.7. Since their seminal studies, empirical observations have become more nuanced. A now often recurring criticism of their observations includes their consideration of cities from distinct and different urban system together. The size efficiencies associated with the increased density in the authors data have been argued to fade if one were to control for continental groupings and fuel prices among other characteristics (Gordon, 2008; Karathodorou, Graham, and Noland, 2010; Steemers, 2003).

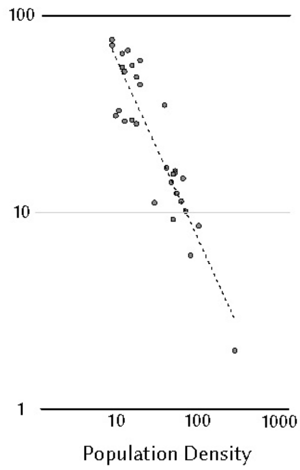


Figure 2.7: Variation in annual transportation energy consumption, GJ, and population density, prs/ha, across major cities during the 1980s – data from Newman and Kenworthy (1999).

More recently, in a case study performed on neighborhood level low- and high-density areas in Toronto, Norman et al. (Norman, MacLean, and Kennedy, 2006) note higher per capita efficiencies associated with the high-density development in transportation, building operations, and material sectors. Similarly, O'Brien et al. (2010) suggest an overall decreasing behavior for the net energy use in cities, consisting mainly of household and transportation uses, with increasing housing density. Studies of this nature, which often indicate that increasing population/built density is correlated with increasing urban efficiencies, are mostly rooted in and can be explained by a theoretical expectation derived from thermodynamic principles regarding consumption and accessibility within more densely built and populated areas. Taking increasing population density to indicate denser construction forms, the more compactly built forms tend to provide smaller surface-to-volume ratios and hence lower potential environmental losses and overall urban consumption (Mohajeri, Gudmundsson, and Scartezzini, 2015). Theoretical modeling of energy demand for different urban morphologies based on four case study cities of London, Paris, Berlin, and Istanbul confirms this by finding potential for significant savings achievable through higher built densities (Rode, Keim, Robazza, Viejo, and Schofield, 2014). What is, however, presently missing from these studies is a consideration of whether these population size/density efficiencies are influenced by the geographic scale at which urban units are aggregated.[8] Nevertheless, the general consensus amongst the energy-focused studies points at two main reasons as to why high density built-environment and cities are expected to be more efficient in their energy use. First, the compactness and higher densities result in higher efficiencies within the building and that reduced time of travel and communication characteristics of higher densities are advantageous towards better transportation performance (Hui, 2001).

[8] Although using deeply flawed methodology, Bettignies et al. (2019) ostensibly attempt to highlight this by examining and comparing population and population density elasticities of energy use for both administratively defined cities of different sizes and their administrative subdivisions.

Outside this energetic-focused literature, the parallel for this influence of density on transport-related efficiencies can be found in the studies of urban travel time. Most generalized, the travel time budget is a notional universal constant that allocates the temporal budget an individual dedicates to daily travel and commute. This would in turn influence their spatial choices for housing and employment among other things.[9] Given the stability of empirical values, when travel time budget is assumed to be constant, the measures reducing journey times are seen as facilitating and encouraging individuals to increase distances traveled.[10] From such a perspective, then, it is the overall density of the urban environment that would help with efficiencies expected in terms of infrastructure provided and energy consumed. Ewing et al. (1994), investigating six communities in Florida, observe a significant effect of density on vehicle hours

[9] This is often empirically estimated as somewhere between 1.1-1.3 $^{\mathrm{hrs}}/_{\mathrm{prs}\cdot\mathrm{day}}$ (Bieber, Massot, and Orfeuil, 1994; Hupkes, 1982; Schafer and Victor, 2000; Vilhelmson, 1999).

[10] As Gunn (1981) points out these, at the time, would have included almost all travel models and the subsequent policies based on travel time minimization principles.

traveled per person, with core urban areas exhibiting almost half as many hours traveled per person as low-density suburbs, suggesting 'density, mixed use, and a central location all appear to depress vehicular travel'. Analyzing five neighborhoods in San Fransisco, Kitamura et al. (1997) note a similar correlation between neighborhood spatial characteristics, e.g. density, accessibility, and land-use mix, and the mode and mix of the trips generated.[11]

[11] The authors do, however, caution that individual attitudes appear to more strongly control travel demand and mode suggesting that planning measures, such as increased density, in the absence of behavioral shifts are unlikely to be successful in delivering anticipated efficiencies (Kitamura, Mokhtarian, and Laidet, 1997).

Choice of spatial scales, however, remains a source of contention. Disaggregated and at an individual level, it is acknowledged that variations exist and both travel time and monetary expenditure tend to exhibit large variations (Kirby, 1981). Mokhtarian and Chen (Mokhtarian and Chen, 2004) review over 24 works pertaining to travel time expenditure investigating the constancy of individuals' travel-time budget. Although some patterns of travel time expenditure partially correlates with other individual and spatial characteristics, e.g. car ownership, employment, household income, spatial structure, and population density, the idea of a constant travel-time budget only holds at highly aggregated scales. This means that both population and its density would be of less significance individually with the time differences only significant when considering urban areas that exhibit low-density patterns and yet house a large population which is then more likely to travel longer distances. The question that arises is twofold. There is first the matter of whether there would exist a size-threshold for the increased efficiencies and productivities that accompany urban size. And, more crucially, over what distances and connectivity patterns such limits might be at work.

2.2.2. Economic arguments and observations

Similar to the energetic arguments, the early 1990's saw a significant resurgence of interest and activity in economics concerned with agglomeration-based perspectives on urban and regional productivity. These works made a very strong case as to why spatial concentration of activities, particularly in and around cities, plays a fundamental role in the growth of the economy (Glaeser, Kallal, Scheinkman, and Shleifer, 1992; P. Krugman, 1991a; Porter, 1990; Scott, 1988). However, starting with the perspectives formulated prior to the 1990s, agglomeration effects and the nature of mechanisms explaining size-related productivities and efficiencies can be considered to fall into three categories.[12] First are the internal scale economies. These are those related to mechanisms by which production efficiencies are achieved by individual firms due to their larger size, rather than that of the city hence 'internal'. Although

[12] Note that these are following common formulations of Ohlin (1933) and Hoover (1937) and particular conceptual, and sometimes subtle, differences with alternative framings of agglomeration and industrial clusters, e.g. that of Marshall (1890) or Chinitz (1961), are beyond the scope of this volume.

these do not concern the cities within which the firm is located, they are still in nature spatial and include the benefits that firms accrue by targeting and intensifying resources and investments at one location in space. The second category, economies of localization, concerns mechanisms that increase firms productivity as a results of clustering of firms belonging to the same industry through input sharing, labor market pooling, and knowledge spillovers (Marshall, 1890). And finally, there are urbanization agglomerations which are assumed to operate based on the same three factors of input sharing, labor market pooling, and knowledge spillovers albeit, in contrast to the previous two, due to urban population size and diversity (Gomez-Lievano, Patterson-Lomba, and Hausmann, 2017; Jacobs, 1970).

A majority of the empirical evidence for agglomeration economies and clustering benefits, from those pertaining to knowledge spillovers to labor-market pooling and input sharing, is concerned with localization economies within different industries while urbanization effects are believed to be industry idiosyncratic as industries balance productivity gains against various congestion penalties (Moomaw, 1981; Rosenthal and Strange, 2003; Sveikauskas, 1975). Based on a probabilistic model by Jovanovic and Rob (1989) that codified knowledge spillovers as the 'diffusion and growth of knowledge' that takes place through local interactions of random individuals who would then further exchange and augment ideas in each given time period, Rauch (1993) hypothesized that increased average human capital, gaged through the proxy of education level and work experience, would further facilitate such growth. In an examination of data pertaining to metropolitan statistical areas (MSA) in the US, Rauch (1993) estimated that addition of a year to the average education level would increase total-factor productivity[13] by 2.8%. The fundamental mechanism in such agglomerative models can be seen as larger local economies, i.e. those with higher levels of activity that would allow for a larger and more diverse intermediate producers which in turn increases the production productivity of the final goods (Ciccone and Hall, 1996).

[13] the portion of total output not explained by capital and labor inputs.

Ciccone and Hall (1996), however, criticized preceding studies for their focus on returns to total size, either urban population or industry employment, believing density more suitable. In their study of productivity variations across US states, Ciccone and Hall (1996) attributed larger-scale, state level, variations in output productivity to differences in employment density at lower spatial scales, county level, whereby a doubling of employment densities in a lower-scale unit causes the average-aggregate productivity in the parent unit to increase by about 6 percentage points. They also confirmed that density is a better descriptor of increasing returns than absolute size would be.[14] More recently, Rosenthal and Strange (2003) consider the effects of geographic distance and scale on the strength of associated agglomeration elasticities associated with the three aforementioned categories. They note that localization effects drop rapidly with increased distance from industry center before more or less stabilizing over longer distances. Using a software industry example, they state that the addition of the same number of new workers to the 1-5 and 5-10 mile rings would create virtually identical industry growth rates. Meanwhile the same addition to the immediate 1 mile radius would enjoy rates 10 times higher. Similar to the energetic evidence, studies tend to highlight the potential inter-play of size and density with the spatial scales over which these distances and connectivity limits emerge remaining subject to ongoing debate (Combes, Duranton, Gobillon, Puga, and Roux, 2012).

[14] We will see in chapter 3 how idealized provisions of connectivity and infrastructural efficiency would inter-link total size and density where congestion costs are only kept in check if larger population sizes area also denser territorially potentially explaining this suitability of density over total size observed much earlier by Ciccone and Hall (1996).

The perspective on urban agglomeration has, since the early 2000s, evolved and shifted towards a more nuanced and complex reading where observations regarding the effects of geography, density, and diversity vary depending on where from in the world the data originates. In places with a substantial split between urban and rural territories, e.g. the US, Canada, and Australia, empirical observations conform closely to these types of urban economic arguments such that bigger cities are broadly more prosperous.[15] These urban hierarchies are often visually explicit and recognizable in these countries where the patterns of urbanized area are reminiscent of those either theoretically derived by Christaller (1966) or Lösch (1954).[16] This simple reading of the evidence on cities, as suggested in chapter 1, broadly paints the issue of urban performance as one of scale whereby *bigger does more with less.*[17]

[15] See the meta-analyses performed by de Groot et al. (2007) and de Melo et al. (2009).

[16] These are seen as series of regional economies with each locality having dominating capital surrounded with satellite cities of second and third tier.

[17] It should be noted that there remains a lack of consensus as to what constitute 'large' or 'medium' when discussing city size.

At a broader system-of-cities level, the addition of more diverse data from across different and unrelated urban systems has also caused few questions to rise as to the significance and meaning underpinning statistical regularities at a system-level.[18] While size-related patterns observed in the context of living organisms' body size and metabolic rate are rooted in the geometry of their physical shapes and thermodynamic/energetic efficiencies, questions remain as to what mechanisms, from an economics perspective, underpin such observations across cities as the physical arguments do not ostensibly connect with economic issues.[19] There remains a sense, however, that these empirical patterns are somewhat related with city productivity and performance despite the potential effects of random stochastic processes in their manifestation. This line of thinking was mostly reinforced by Fujita et al. (1999) who showcased a model within a NEG framework that derived and resulted in a distance-based urban hierarchy and concluded that such regularities cannot entirely be a

[18] Refer back to section 2.1.4 for details.

[19] Zipfian rank-size distributions can, however, be derived under various assumptions and there are numerous attempts at replicating them, see Gabaix (1999) and Fujita et al. (1999).

result of random processes and require underlying behavioral principles that capture how people function and interact, Figure 2.8.[20] Although the urban growth process within such frameworks is conceptualized as a balancing act between increasing productivities and escalating congestion costs, the formulation and consideration of the congestion related penalties remain mostly abstract (Abel, Dey, and Gabe, 2012; Henderson, 1975). Another remaining question concerns the occasional empirical non-compliance of large cities at the top of these urban hierarchies with these statistical patterns that are routinely observed to more strongly and closely match systems with larger number of medium and small cities (Arcaute et al., 2015; Cottineau, 2016).

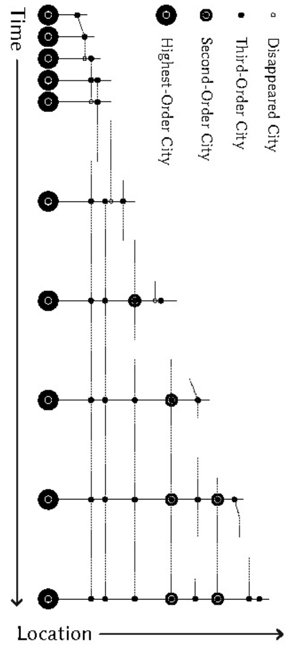


Figure 2.8: Schematic showing the distance-order hierarchy for an evolving urban system over time with discontinuous vertical lines tracing cities from their creation to disappearance across time steps that are not shown – adapted from Fujita et al. (1999).

[20] See also Henderson and Venables (2009) for a more recent example.

Another side of the *bigger does more with less* consists of arguments relating to the higher chances of diversity inherent to larger places. These stem from the dichotomy of specialization-diversification where the former is seen as hindering and the latter helping economic growth (Jacobs, 1970). Similar to the urban hierarchy and rank-size observations, the evidence in this space appears to depend on how measurements are taken and analyzed (Henderson, Kuncoro, and Turner, 1995). The meta-analysis of the evidence suggests a lack of statistically significant advantage in either specialization or diversification in general with the advantages in specialization and diversification depending on industries' state of maturity in a given city (Beaudry and Schiffauerova, 2009). Hence, while the early narrative portrayed and evidence gathered throughout the 1990's follow Glaeser's (1992; Jacobs, 1970) arguments that the urban hierarchy includes fewer larger more diversified and productive cities at its top followed by an increasing number of smaller more specialized places with decreasing productivities, the current set of evidence suggests that in many parts of the industrialized world the picture is more varied country by country. Finally, a broad review of theoretical works and models on agglomeration points to half a dozen mechanisms that would generate and explain patterns of spatial concentration of activities and hence size-related urban productivities and efficiencies (Duranton and Puga, 2004; Puga, 2010). Common across all these mechanisms is the underlying notion that a concentrated *larger* unit would, statistically, provide better chances of:

1. more efficiently sharing common infrastructure and resources,
2. more reliably matching complementary needs, and 9
3. facilitating learning through increased interactions.[21]

[21] Puga (2010), however, does note an ongoing difficulty in testing these mechanisms empirically as they all provide for a similar outcome, i.e. increased productivities and efficiencies with increased urban size, leaving the disaggregation and attribution of the magnitude of individual mechanisms' contribution challenging.

2.2.3. Connectivity arguments and observations

The implicit reliance of the mechanisms noted in the previous section on the connectivity intuitively assumed to be better afforded in larger cities, brings us to the arguments and observations of the effects and influence of connectivity, and particularly that of transport and infrastructure, on urban productivity and its patterns.

Analytical models of the New Economic Geography (Fujita, Krugman, and Venables, 1999; Puga and Venables, 1997) are quite insightful as to the impact provision and cost of transport have on the urban productivity and its geographic patterns. This family of theoretical models can be quite detailed and particular as to the differences in impacts of different types and configurations of connectivity. Studies have considered connectivity patterns contrasting those that specifically provide inter-regional transport through a common hub unit (Fujita and Mori, 1996; Puga and Venables, 1997). Similarly, they have examined the implications of different 'types' of roads distinguishing between those of long- or short-distance nature abstracting inter- and intra- regional connectivity (Fujita and Thisse, 2002; Ottaviano, 2008; Puga, 2002). Broadly, the theoretical expectation of these models suggests that increased accessibility would promote agglomeration although in an asymmetric manner (Rodríguez-Pose, Crescenzi, and Di Cataldo, 2018). This is to say that an increased inter-regional connectivity, i.e. the hun-spoke connectivity or the long-distance roads, would intensify agglomeration within the hub region at the expense of the other connected regions. This is in contrast with the short-distance connections which are thought to effect the regions in which they are provided more positively (Ottaviano, 2008).

The empirical evidence more or less supports these expectations. In examining the effects of the provision of inter-state highways on the economic activity across the US, Chandra and Thompson (2000) noted a positive economic effect in the counties adjacent to the highways but only to the detriment of non-adjacent counties in the states with an overall net zero effect. In a similar fashion, an examination of the influence of transport infrastructure on the economic performance of the European regions found very little evidence of positive impacts of road infrastructure endowment once other social characteristics where accounted for (Crescenzi and Rodríguez-Pose, 2012). Meanwhile, Del Bo and Florio (2008) in investigating the effects of infrastructure capital on income level and growth, disaggregated by infrastructure type, at NUTS2 levels among EU27 member states report positive correlations especially for information followed by overall accessibility infrastructure.[22] Similarly, Caragliu et al. (2011) in their review of the characteristics of *smart cities* observe significant positive correlations between both multi-modal accessibility and length of public transport network and per capita gross domestic product when considering the partial correlations among 6 smart city indicators for over 200 European contiguous urbanized area (UA) areas between 2003-2006. While they do caution that the causality is likely to be bi-directional in case of transport provisions, density, and urban output, Caragliu et al. (2011) posit that the strong correlations suggest overall density of public transport networks and multi-modal access alleviates potential congestion effects that the overall built density of urban areas might cause.

[22] It should, however, be noted that when considering transport infrastructure their proxy data is only focused around density of the routes, e.g. km/area of motorways or rail lines, with congestions proxied through inter-regional truck traffic.

The use and specification of New Economic Geography models at an urban system level faces few challenges, however. Despite the detailed and particular insights they provide, the theoretical models of this family can very quickly become both analytically and computationally intractable without stylistic simplifications such abstracting physical features, e.g. urban size, infrastructure, etc., in form of their effect on utility functions or limiting the number of interacting regions or the industries within them (P. Krugman, 1991b). Consequently, the notion of inter- and intra-region connectivity explored by these models are abstracted to the units definition within the model exacerbating the difficulty in identifying the correct spatial scales for empirical examination of these effects. These are evident in observations that physical transport networks are often arbitrarily truncated at regional or national boundaries in empirical studies potentially masking network-level effects and leaving the effects of geographic scales themselves unexplored (Holl, 2006; Laird, Nellthorp, and Mackie, 2005). The literature analyzing connectivity-driven agglomeration efforts, whether arguing for or against, is then often locked on an arbitrary designation of regional and metropolitan spatial scale constrained by geographies over which data is available (Combes et al., 2012; Glaeser and Gottlieb, 2009; P. Krugman, 1995; Overman, Gibbons, and Tucci, 2009).

2.3.
Summary of research needs

In this chapter, we first outlined a broad bibliometric overview of the disciplines dedicated to studying cities. This we followed with a more focused review of the evidence underlying a bigger-does-more-with-less view of urban areas in regard to their population size and its spatial organization. This review of urban agglomeration and productivity highlighted a recurring call for the examination of their combined effects with that of geographic scales on agglomeration forces (Holl, 2006; Laird et al., 2005; Puga, 2010) and references to the opposing effects of inter- and intra-regional connectivity (Fujita and Mori, 1996; Fujita and Thisse, 2002; Puga, 2002). Meanwhile, our initial bibliometric overview enabled us to identify the emerging body of literature, i.e. the complexity, allometrics, and the new science of cities, that provides a connective tissue bringing together various disciplinary interests in cities from urban economics and economic geography to industrial ecology and urban metabolism.

Urban scaling frameworks also provide a number of additional advantages when compared with the approaches of the existing literature. These scaling perspectives try to capture a system-wide formulation of cities and urban networks which brings about a fresh framing and line of argumentation to the issues surrounding population density, connectivity, and productivity. The NEG type frameworks pioneered by Krugman (1991b), while powerful, also operate on a particularly specific set of assumption about the urban system and its behavior. The scaling-type arguments are however more inherently general in their choice of assumptions which the wider scholarly audience, particularly transportation and behavioral scientists, would regard as rather more palatable and reflective of people's behavior. Having been developed at the interface of disciplines with a broader view of cities, scaling arguments and frameworks stand more accessible to people from different disciplines including urban economists (Florida, Adler, and Mellander, 2017; Glaeser, Ponzetto, and Zou, 2016; Miguélez and Moreno, 2013). Consequently, the broad appeal and utility of allometric models, as we will see, lie in their generalization of system behavior across different sizes and spatial scales (Bettencourt and West, 2010; West, Brown, and Enquist, 1997). Complemented by the increasing availability of data and the rise of geographic information system enabling actual measurement and collection of empirical observations, these allow for a practical examination of the roles of geographic scales and population density with respect to urban productivity.

This brings us back to our overarching research questions previewed in the previous chapter and the means by which we intend to approach them. Our broad research theme targets the effects of geographic scales as one of the main open lines of inquiry. Examining the application and suitability of the emerging allometric arguments in studying these scale effects, our questions can be further elaborated and disaggregated as:

A. do allometric arguments regarding connectivity and agglomeration remain valid across different scales and definitions of

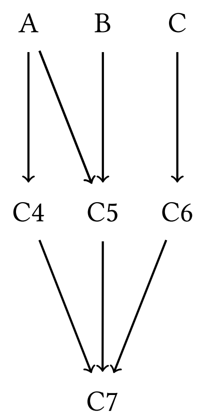


Figure 2.9: Schematic of the connection between chapters and research questions.

city boundary?

B. to what extent are such scaling insights transferable between different cities and systems of cities? and finally,
C. what are the effects of spatial scales and distance on the geographical patterns of connectivity-based agglomeration within such scaling frameworks?

In this work, we therefore investigate both the influence of spatial scales on observations of urban agglomeration effects and the extent to which the application and interpretations of the emerging family of scaling frameworks can be operational and transferable from a practical point of view, Figure 2.9. In doing so, the rest of this volume provides one of the first studies operationalizing existing allometric urban models, specifically that of Bettencourt's (2013), for a multi-scale exploration of urban networks and their hierarchical configurations and obtaining practical planning perspectives. The volume also provides one of the very first spatially multi-scale empirical assessment of urban networks which incorporates various different regional boundaries. Using such scaling frameworks in a sense encapsulate a majority of the existing urban economic debate while enabling a shift in perspective and adaptability to other urban-oriented studies. As mentioned, this allows for a more system-wide examination which is simultaneously more accessible and approachable to a broader inter-disciplinary audience. Finally, as we will see in this volume, while a scaling approach does not yet provide all the answers, it offers crucial flexibility for multi-scale analyses and gives surprising insights vis-à-vis interactions of density and connectivity in the context of the overall urban performance balance.

3

Methods, Data, and Approach

Material prepared for this chapter have been used in the following–

Arbabi, H., Mayfield, M., & McCann, P. (In Press). Productivity, Infrastructure, and Urban Density—an Allometric Comparison of Three European City-Regions across Scales. *J R Stat Soc A Stat*.

Arbabi, H., Mayfield, M., & McCann, P. (In Press). On the Development Logic of City-Regions: Inter- Versus Intra-City Mobility in England and Wales. *Spatial Economic Analysis*.

Arbabi, H., Mayfield, M., & Dabinett, G. (2019). Urban Performance at Different Boundaries in England and Wales through the Settlement Scaling Theory. *Reg Stud*, 53(6), 887-899.

In the past decade, with growing abilities to collect, share and analyze larger bodies of data pertaining to urban settlements, an understanding of cities and their properties as population scaling functions, formulated in the vein of similar allometric relations underlying the growth and size of organisms, has gained more traction both analytically and empirically (Bettencourt, Lobo, Helbing, Kühnert, and West, 2007). More importantly, this allometric line of thinking has already made an impression on the planning and economics literature. While part of this influence has been implicit in the form of concurrent observations of city rank–size distributions (Cheshire, 1999), others like Glaeser have been more explicit in the use of and reference to this field of literature, its theoretical frameworks and the models it provides in their own works (Glaeser et al., 2016).

3.1. A unified theory of urban living

Bettencourt and West (Bettencourt and West, 2010), supported through a series of analyses investigating large American, Chinese, and European urban datasets (Bettencourt et al., 2007), introduce the notion of 'universal features'. Their framework puts emphasis on the size of a city, often considered as the city's aggregate inhabitant population, as the primary driver of the average-aggregate urban characteristics further complemented with the city's geography, physical design, and history. They formalize these correlations of urban properties with city size as:

$$F(N) = F_0 N^{\beta} \tag{3.1}$$

or more conveniently linearized through a log-log transformation

$$\ln F(N) = \ln F_0 + \beta \ln N \tag{3.2}$$

where $F(N)$ denotes the average-aggregate urban characteristic of choice for population size N, e.g., gross value-added (GVA), urbanized land area, employment etc., F_0 is the baseline prevalence of F, and β is the exponent determining the nature of the scaling relation.[1] To explain these observations, a variety of urban models have been developed that yield such scaling behaviors for aggregated average response of urban attributes. These include models rooted in probabilistic conceptualizations of activities taking place in cities and the portion of the population contributing to them (Gomez-Lievano et al., 2017) and those which are based on network realizations of the interactions between inhabitants and/or the geographical embedding of such networks within cities (Yakubo, Saijo, and Korošak,

[1] Baseline prevalence of various phenomena, F_0, in essence captures the system-wide representative 'per capita' expectation of a given phenomenon excluding the scale effects that are characteristic of the variable growth rate of allometric relationships.

2014). More importantly, these studies also make the observation that certain properties consistently exhibit specific scaling regimes with metrics describing built infrastructure showing sub-linear scaling, demonstrative of increasing efficiencies in larger cities, and those descriptive of individual interactions and processes, i.e. wealth, information and innovation, etc., displaying a super-linear scaling[2] (Bettencourt, Lobo, Strumsky, and West, 2010; Schläpfer et al., 2014), Figure 3.1. Similar allometric behaviors across different countries and for various other indicators have been investigated including but not limited to those investigating the city total area (Longley, Batty, and Shepherd, 1991; Nordbeck, 1971; Paulsen, 2012), length and area of infrastructure, e.g. length of road networks (Y. Chen, 2010; Masucci, Arcaute, Hatna, Stanilov, and Batty, 2015; Mohajeri and Gudmundsson, 2014; Mohajeri, Gudmundsson, and French, 2015), electricity cables (Bettencourt et al., 2007; Kühnert, Helbing, and West, 2006), etc. In fact, research from the other communities, specifically the industrial ecology/urban metabolism cluster, has also moved towards incorporating a number of ideas found predominantly within these physically-inclined publications. Bristow and Kennedy (2012, 2013, 2015) through a series of studies have tried to move from the traditional equilibrium-based MFA to a non-equilibrium understanding of cities as entropy-maximizing machines destroyers of exergy. These have included efforts to seek thermodynamic clues in Zipfian distributions of cities and their efficiency in energy consumption compared with that of other distributions (Bristow and Kennedy, 2013) and studying energy requirements of Hong Kong and Singapore to sustain their economies under scaling-based future growth scenarios (Bristow and Kennedy, 2012). Others have also explored the existence of energy-themed universal features and the application of entropy-based analogies for urban CO_2 emissions and energy dissipation (Fragkias, Lobo, Strumsky, and Seto, 2013; Horta-Bernús and Rosas-Casals, 2015; Mohajeri, Gudmundsson, and French, 2015; Mohajeri, Gudmundsson, and Scartezzini, 2015). The universality of these scaling exponents, on the other hand, has been questioned and their sensitivity to the choice of settlement boundary has recently been pointed out (Arcaute et al., 2015; Cottineau, Finance, Hatna, Arcaute, and Batty, 2018), especially in scaling patterns relating to energy, emissions, and innovation where different studies report exponents with broad or conflicting interval ranges.

[2] Such a super-linear scaling where a doubling of population would result in a, say, 16% increase of the indicator's growth rate, e.g. per capita GVA, provides similar incentives to those of the economic agglomeration in desiring larger cities.

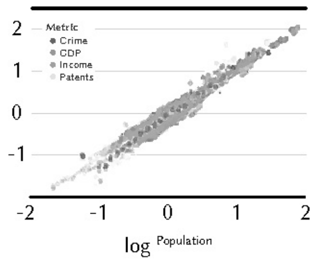


Figure 3.1: Data from some 360 US metropolitan areas showing scaling between average-normalized city metrics and population on a double logarithmic plot – adapted from Bettencourt and West (2010).

More relevant to our interests in this work, these frameworks formalize agglomeration effects captured within a system of cities and hence provide a means to evaluate idealized counterparts to cities of a given population size. What can be taken as an idealized city is then derived from a power-law scaling regime that underpins an overall urban system to which a given set of cities belong. These are hence frameworks of a system of cities based on the relationship between agglomeration forces and the costs of human interactions. On the basis of these, from an agglomeration-based scaling point of view, cities would follow sub- and super-linear population scaling for infrastructure, i.e. length of road network, total urbanized area, etc., and economic output respectively with the magnitude of these elasticities, here the scaling exponent, a function of geographic geometry and mobility. In this context, the idealized counterpart to a city, not an intrinsically ideal city, would be that which shows the least deviation from the desired productivity and efficiency elasticities for the same population size.

3.2. Social reactor model

In this section, we briefly outline and describe the social reactor model as formulated by Bettencourt (2013). This volume will make use of this model and its application without any adjustments in or improvements to its original formulation by Bettencourt. In setting up an idealized scaling model of cities, Bettencourt (2013) starts from four simple assumptions:[3]

1. the average aggregate socio-economic product is a linear function of the sum of all local interactions (Jones, 2017),
2. urban population is mixing uniformly and that each individual has the minimum resources needed to fully travel and experience the city (Glaeser and Kohlhase, 2003),
3. individual baseline production is invariant of city size (Szüle, Kondor, Dobos, Csabai, and Vattay, 2014),
4. and finally, the urban infrastructure is embedded as a hierarchical network that keeps all individuals connected through its incremental and decentralized growth (Samaniego and Moses, 2008).[4]

The model also parametrizes and expresses the geometry of the city and the average inhabitant travel path through their Hausdorff fractal dimensions, D and H respectively.[5] Out of the four, the first assumption can be formalized as

$$Y = \bar{g} a_0 l \frac{N^2}{A_n} \tag{3.3}$$

where Y is the average economic output, $\frac{N^2}{A_n}$ the density of the upper limit of total encounters possible – $N \times (N-1) \approx N^2$ for large populations – over the urbanized area A_n, $a_0 l$ the average effective interaction cross-section and travel path of an individual respectively

[3] Note that the citations against each assumption point towards studies that provide empirical evidence in support of Bettencourt's assumptions and are not necessarily those cited by Bettencourt (2013).

[4] Although the first two assumptions may appear contentious, it should be noted that the first is supported by current empirical observations and generally agreed upon across other urban scaling models (Gomez-Lievano et al., 2017; Sim, Yaliraki, Barahona, and Stumpf, 2015; Yakubo et al., 2014) while the second is ultimately an idealized and stylized assumption that affects the value of the scaling exponent and not the existence of an overall population power-law relationship.

[5] In brief, Hausdorff number provides a generalized indication of roughness so that for smooth features such as a Cartesian point, straight line, or plane, Hausdorff dimension coincides with the topological dimensions 0, 1, and 2, respectively, while for rough fractal geometries Hausdorff dimension would exceed the topological dimension (Mandelbrot, 2004).

and hence the average effective area, and $\bar{g}$ the average output of each encounter. The product $\bar{g}a_0 l$, hereafter referred to as G, describes the baseline human production indicated in the third assumption and embodies the average sum-total of individual output independent of population size ($\frac{dG}{dN} \approx 0$). The second assumption then derives a generic scaling of city area by equating per capita mobility costs, i.e. cost of travel, and per capita economic output, i.e. minimum resources needed for travel

$$G\frac{N}{A} = \epsilon A^{\frac{H}{D}} \tag{3.4}$$

where the left-hand side is the absolute minimum outcome available to each individual across its path in the city and the right-hand side equals the cost of the associated travel with ϵ the per person per unit length cost of travel, and $A^{\frac{H}{D}}$ the travel path length as generalized in terms of path's Hausdorff dimension, H, and the topological dimension of the city, D.[6] Re-arranging Equation 3.4 gives

$$A = aN^{\alpha} \tag{3.5}$$

where $a = (\frac{G}{\epsilon})^{\alpha}$ denotes the baseline area, and $\alpha = D/(D + H)$.

Bettencourt then writes the hierarchical infrastructure assumption as

$$N_i = b^i, \ \forall i \mid 0 \leq i \leq h \tag{3.6}$$

where b is the ratio of infrastructure units between successive layers, e.g. alleys to roads, or roads to motorways, and N_i is the number of units at level i of the infrastructure hierarchy, with $i = 0$ denoting the largest layer and $i = h$ the smallest such that the number of smallest infrastructure units $N_h = b^h$ equals the total population N. Infrastructure geometry can then be formalized at each hierarchical level as

$$\begin{cases} a_i = a^{\alpha}(b^i)^{\alpha-1} \\ l_i = \dfrac{a_i}{l} \\ s_i = s_*(b^{h-i})^{1-\frac{H}{D(D+H)}} \end{cases} \tag{3.7}$$

with a_i, l_i, and s_i denoting the land area an infrastructural unit belonging to level i crosses, its length, and its transverse dimension, respectively and l the average separation of city blocks. It is easy to see that the scaling regime of unit area a_i follows that of Equation 3.5 ($\frac{A}{N} \propto N^{\alpha-1}$) and that $l_i = \frac{a_i}{l}$ is due to the area-filling nature of an infrastructure that would grow incrementally to keep all inhabitants connected. Since the infrastructure network is to

[6] Note that here we are using A as opposed to A_n which denotes the total and not the urbanized network area.

keep all inhabitants connected, Bettencourt takes the average distance between individuals to equal average infrastructure length per capita and as such similar to a_i, the scaling regime governing the transverse dimension s_i follows from Equation 3.5 but as a length ($N(\frac{A}{N})^{\frac{1}{D}} \propto N^{1-\delta}, \delta = \frac{H}{D(D+H)}$). This final scaling also contains the assumption that the base transverse dimension of the smallest infrastructural units, s_*, do not scale with population and are constant across cities. While abstract, the assumption follows the intuitive observation that building blocks of build infrastructure, e.g. doorways and taps, show common cross-sectional size independent of cities. With these, the total urbanized area can be written as

$$A_n = \sum_{i=0}^{h} s_i l_i N_i = A_{n0} N^{1-\delta} \tag{3.8}$$

where $A_{n0} = \dfrac{s_* a^\alpha}{l(1 - b^{\alpha+\delta-1})}$.

Bettencourt then adopts a parallel circuit model with the assumptions that the total current, J, is conserved across infrastructural level, i, and that similar to s_*, current density and inhabitant velocity over smallest infrastructural units are population invariant leading to $J = J_h = J_* N$. Taking the resistance of each infrastructural hierarchy as a function of its geometry with the units in parallel, the combined resistance at each hierarchy can be written as

$$R_i = (\sum_{i=1}^{N_i} \tfrac{1}{r_i})^{-1} = \tfrac{r_i}{N_i} = r \tfrac{l_i}{s_i N_i} \tag{3.9}$$

where r is the unit resistance per unit length and transverse area. Finally, the total power dissipated in mobilizing over the infrastructural network can be formalized as

$$\begin{aligned} W &= \sum_{i=0}^{h} W_i = \sum_{i=0}^{h} {J_i}^2 R_i = J^2 \sum_{i=0}^{h} R_i \\ &= W_0 N^{1+\delta} \end{aligned} \tag{3.10}$$

where $W_0 = \dfrac{a^\alpha r {J_*}^2}{l s_* (1 - b^{\alpha+\delta-1})}$.

Put together, the four assumptions result in

$$\begin{cases} Y(N) = Y_0 N^{1+\frac{H}{D(D+H)}} \\ W(N) = W_0 N^{1+\frac{H}{D(D+H)}} \\ A_n(N) = A_{n0} N^{1-\frac{H}{D(D+H)}} \end{cases} \tag{3.11}$$

where Y, A_n, and W are the expected average-aggregate economic output, urbanized area, and mobility costs respectively, Y_0, A_{n0}, W_0 the baseline prevalence of Y, A_n, and W all functions of G, N the population size, and the exponents $\beta = 1 \pm \frac{H}{D(D+H)}$ functions of the geometry of the city, D, and the geometry of the average individual path, H, in effect characterizing mobility and accessibility within the city. Most empirical studies also observe the statistics of such fluctuations to be Gaussian and zero-centred for the log-transformed Equation 3.2 for a range of urban indicators (Bettencourt et al., 2007; Gomez-Lievano, Youn, and Bettencourt, 2012). Note that while Bettencourt originally derives the scaling of W by modeling the built-up infrastructure as a fully parallel hierarchical network of resistors, changing the configuration of the resistance model used would only affect the magnitude and composition of the pre-factor, W_0, and not the general scaling developed for average-aggregate W.[7]

[7] Also, note that the formulations in Equation 3.11 represent the average expected values describing the urban behavior across an entire urban network. For the formulation to be exact the inclusion of a fluctuation term, ξ, is required (Bettencourt and Lobo, 2016).

3.2.1. Urban mobility and a size-cost performance balance

Imposing only real-life geometric constraints puts the fractal dimension of the city, D, somewhere in the range $[2, 3)$. Similar considerations would result in the Hausdorff dimension of the travel path, H, to be confined to $[0, D)$ resulting in a range of $[0, \frac{1}{4})$ for $\frac{H}{D(D+H)}$. As such, in agreement with agglomeration theory, the model expects increasing output productivities and infrastructural efficiencies for larger cities, i.e. a super-linear scaling of Y and sub-linear scaling of A_n whereby per capita wealth generation in cities increases with population size while per capita need for infrastructure decreases. In developing a theoretical and idealized approximation of urban networks, city geometry can be taken to be 2-dimensional topologically, $D = 2$, while Bettencourt's second assumption regarding full accessibility of the city implies a fully linear average travel path, $H = 1$. Consequently, Bettencourt's theoretical expectation of ideal urban networks is comprised of a super-linear scaling for economic output with the exponent $\beta_Y = \frac{7}{6}$ and a sub-linear scaling of urbanized area with the exponent $\beta_{A_n} = \frac{5}{6}$ in agreement with most empirical observations for various urban networks in the United States, East Asia, and continental Europe (Bettencourt, 2013; Bettencourt and Lobo, 2016; Bettencourt et al., 2007). It is useful to emphasize that it is these basic behavioral assumptions about people's interactions and city geometry that gives rise to the theoretical expectations of the exponents of $\beta_Y = \frac{7}{6}$ and $\beta_{A_n} = \frac{5}{6}$. These theoretical exponent values are not assumptions within the model and are functions of the values of D and H with the physical limit of $H \leq D$. An absolute rational upper bound of $\beta_Y = \frac{5}{4}$ can also be assumed to occur at $H = D = 2$ although this would very unrealistically imply that inhabitants on average cover the entirety of the city area routinely. The aggregated evidence across the European countries for OECD's harmonized functional urban areas and the American MSAs does in fact provide a fair match with the theoretical exponents expected at $D = 2$ and $H = 1$, with the latter having provided the dataset on which Bettencourt's model has been based (Bettencourt, 2013; Bettencourt and Lobo, 2016; Bettencourt et al., 2007). Furthermore, since these elasticities, $\beta = 1 \pm \frac{H}{D(D+H)}$, are increasing functions of H, lack of appropriate mobility and access, $H < 1$, whereby individuals' access is limited and constrained to disconnected patches within the city, diminishes super- and sub-linear effects resulting in close to linear relations between economic output and population.[8] What is worth emphasizing before moving on is the average-aggregate systems perspective inherent to the framework. Prefactors Y_0, W_0, and A_{n0} are derived parametrically for the average-aggregate size-scaling of a given number of cities meaningfully belonging to an urban network, say all American cities or all English cities, and given only a single city, there would not then exist a theoretical expectation at a moment in time as no population-related elasticities could be observed given a single data point. There can, however, be a temporal size scaling detailing the growth of the city through time and agglomeration efficiencies compared to the past versions of the city itself (Bettencourt et al., 2007).

Finally, the urban size-cost balance can be formalized in the social reactor model as the economic output less the mobility costs of its generation, $Y - W$. As both Y and W are functions of the baseline human production G, the size-cost balance becomes an optimization exercise with regard to the value of G, Figure 3.2.[9]

As can be seen, size-cost balance grows for increasing values of G in range $[0, G^*]$ reaching its maximum at G^*. However, for increasing values of human production beyond G^* the cost-size balance shrinks resulting in increasingly unstable cities as the costs associated with the mobility processes overwhelm the economic success of the city. This implies that for $G > G_{max}$ the city would break down to smaller functional urban zones. It should be pointed out that it is unlikely to observe a dramatic and/or sudden 'break-down' of an urban area where G exceeds that of the maximum. It is more likely that as G grows for a city within a certain boundary, the city experiences an internal partitioning with regard to its patterns of activity and commuting resulting in units that can be delineated within the original city each with their own closer-to-optimal $Y - W$ balance. An example of this can be seen to some extent in the relationship between London and Heathrow where although the latter comprises part of the contiguous London area, it exhibits self-contained commuting

[8] Figure 3.3 shows example geometries of various Hausdorff dimension H providing a visual shorthand for the comparison between connectivity patterns associated with $H < 1$ and $H > 1$ – note the geometrically discontinuous nature of the Cantor set ($H < 1$) as compared with the Koch curve ($H > 1$).

[9] Beware that the schematic curve included is meant to capture the general form and curvature of the $Y - W$ function and exact gradients of the function before and after G^* depend on the values of D and H among other internal model parameters.

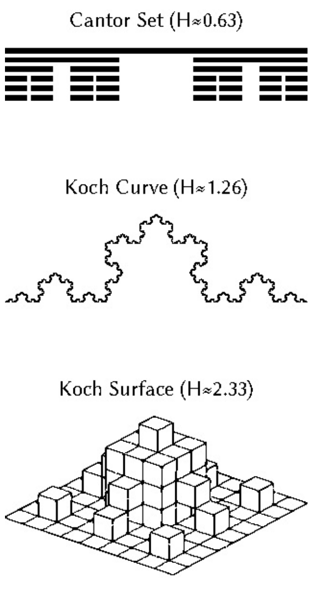


Figure 3.3: Schematic showing fractal geometries and their Hausdorff dimension, from top to bottom, Cantor set ($0 < H < 1$), Koch curve ($1 < H < 2$), and quadratic Koch surface ($2 < H < 3$).

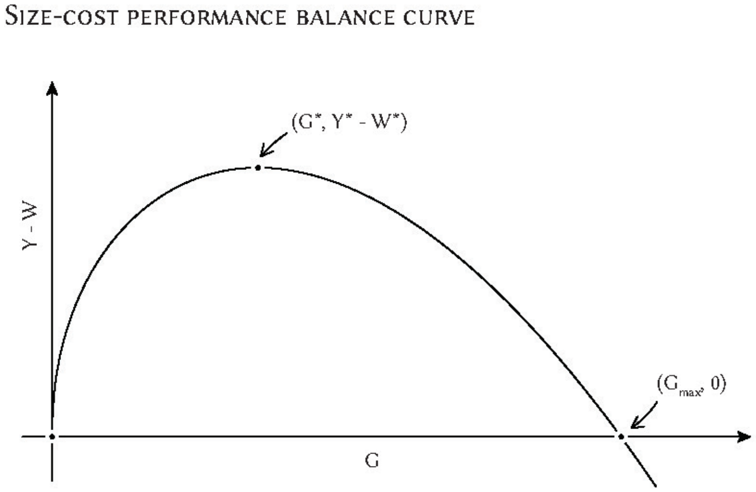


Figure 3.2: Schematic illustration of cost-size balance, $Y - W$, as a function of the baseline human production G.

patterns and is considered a separate functional urban area (Coombes and Office for National Statistics, 2015).

Bettencourt posits that, given an urban network with relatively large number of cities, one would expect to find the statistics of G estimated for all cities therein to hover close to G^* as cities strive to maintain an optimal cost-size balance.[10] Additionally, referring back to the comprising elements within $G(\equiv \bar{g}a_0 l)$, the model provides categorical solutions for cities where the cost-size balance deviates from the optimum. Where $G < G^*$, cities fall short of their economic potential which can be addressed through interventions that seek to increase the effective $a_0 l$, i.e. improvements to mobility and accessibility within cities, enabling more urban interactions and hence higher economic outputs. In contrast, for cities where the economic success of the city has resulted in larger-than-desired growths, $G > G^*$, densification of the built area provides a strategy that would maintain the number of urban interactions and reduce travel paths and hence associated mobility costs concurrently.[11]

[10] Empirical demonstrations of this for the American urban network can be found in (Bettencourt, 2013).

[11] For the sake of completeness, we should also mention that in strictest terms one could also address $G > G^*$ by decoupling the mobility costs, W, from the economic output, Y, through, say, increased energy efficiency of transport systems or more ambitiously by a mass conversion to cycling/walking, such that mobility costs become only nominal.

3.2.2. *Connectivity and the abstraction of mobility and accessibility*

A crucial point to discuss is the interpretations of connectivity, mobility, and accessibility afforded by Bettencourt's model. Within the Social Reactor Model, Bettencourt only considers costs of 'mobility' through the multiplication of ϵ, a per unit length toll for mobility paid by an individual, and the overall distance traveled on average by the individual, $A^{\frac{H}{D}}$, Equation 3.4. The assumption on homogeneous and uniform mixing of the urban population is then articulated by equating this mobility cost, $\epsilon A^{\frac{H}{D}}$ with a lower limit of per capita interactions. In this formulation, given the minimally specified nature of the costs, ϵ, the model's assessment of 'mobility' gages not only the physical characteristics of transport infrastructure, e.g. number and geometry of roads building up of the urban network, but is also implicitly inclusive of the ease by which social interactions take place through various means, their relative affordability. In this sense, the model provides an abstracted and aggregated view of both mobility and accessibility especially when interpreting interventions to restore $Y - W$ balance with respect to G. Because of this, with regard to interchangeable nature of references to transport infrastructure, mobility infrastructure, and accessibility from the model's perspective, we will use the more general notion of 'connectivity' when discussing improvements that would both facilitate an increase in inhabitants' interactions and a reduction of the costs associated with processes enabling them.

3.2.3. Answering our research questions

In chapter 1, we formulated three over-arching questions, the first two of which were:

A. do allometric arguments regarding connectivity and agglomeration remain valid across different scales and definitions of city boundary? and
B. to what extent are such scaling insights transferable between different cities and systems of cities? and finally,

The social reactor model provides a homogeneous framework for addressing and investigating both areas. The existence, stability, and transferability of agglomerative size-productivities can be investigated by examining the exponents of the population scaling of economic output and urbanized area for various urban boundary definitions of different urban systems and across a spectrum of spatial scales. Meanwhile, the relative position of a city unit's baseline productivity, G, relative to the optimal G^*, would offer a glimpse to the potential factors behind their sub-optimal performance providing means to gage nuances that exist across various spatial scales.

It is worth mentioning here that location choices and related arguments are embedded and manifest in the organization of urban systems as the overall urban network would have constituting places of different kinds with different types of interactions. It is implicit within Bettencourt's model that people would have different location choices and are not fixed in place such that location choices between cities affects the system-wide adjustment from an average-aggregated perspective as individual cities grow and shrink in size in response to these choices. Bettencourt's and West's framework (2010) expects cities belonging to a coherent urban network to share and exhibit similar characteristic performance parameters from an average-aggregate perspective. In this manner, these implicit system-wide location choices and evolutionary progress of individual cities can be seen in the complementarity of the scaling exponents of Y and A_n. When cities in a given urban system systematically under-perform economically, they exhibit an expansion of overall urbanized area in order to maintain overall optimality of $Y - W$.[12] This results in cities compensating for smaller than theoretically expected output, at $H = 1$ and $D = 2$, through larger catchment areas.

[12] Note that there are two issues at play here – the overall output or urbanized area scaling exponent close to the theoretical (system-wide performance) and the optimality of the $Y - W$ balance for each city (unit under- or over-performance). These two are to some extent independent as the $Y - W$ trade-off and maximization exists for all urban systems regardless of the magnitude of the agglomeration elasticities caused by particular values of H and D.

3.3. Data preparation and assembly

The rest of this chapter details how we assemble and construct the urban system of England and Wales, and also those of the the Netherlands and Germany for the work presented in chapters 4 to 6, across various spatial scales and the datasets used in estimating values for economic output, Y, and urbanized area, A_n, for each city unit within these differently-scaled realizations of the England and Wales urban network.

3.3.1. Boundaries

The theoretical model and framework discussed previously set out an intuitive and empirically-backed portrait of urban networks. They do, however, remain unclear regarding appropriate delineation of boundaries for what constitute an urban unit, aside from the idealized condition regarding the homogeneous mixing of inhabitants, and as such potentially too sensitive to changes resulting from such boundary differences (Bettencourt, Lobo, and Youn, 2013).

Determining what does or does not constitute a *city* is a task that is both complex and historically divisive. The same is true for finding consensus on fundamental qualities that influence such determinations. In contrast, most individuals share to some extent a common notion of what they may think of as an urban city[13] (Thomlinson, 1969), with a diverse set of methodologies in use to detect urbanized agglomerations based on population density cut-offs (Rozenfeld, Rybski, Gabaix, and Makse, 2011), commuting patterns (Arcaute et al., 2016), and street network characteristics (Masucci et al., 2015). Although most practices use a mixture of criteria combining population

[13] This is perhaps most true in the case of the city in which the individuals themselves reside.

size cut-offs and commuting patterns to define urban boundaries, as these are readily available and relatively easy to keep track of (Cottineau, Hatna, Arcaute, and Batty, 2017), no particularly meaningful and universal cut-off value is available for denoting cities and practice appears to vary across different countries and continents (OECD, 2012, ch. 1).

These variations in definitions pose an issue for the examination of scaling patterns and the generality and universality of their exponents within the range predicted by the SRM (Bettencourt and Lobo, 2016). A number of studies have recently shown wide variations of the expected exponents for different permutations of the urban network where different algorithms or thresholds are used to define cities (Arcaute et al., 2015; Cottineau et al., 2018; European Environment Agency, 2000; NERC Environmental Information Data Centre, 2016). Additionally, in the particular case of England and Wales, the position of London as a first-tier city combined with irregular specialization of some others, e.g. Cambridge, have also been suggested to dis-proportionally affect the consistency of these scaling pattern and the social and economic performance of the urban network as perceived by Bettencourt's model where averaged-aggregate properties of the city take precedent over the nuances of social and economic dynamics at highly disaggregated spatial scales (Arcaute et al., 2015; Cottineau, Hatna, Arcaute, and Batty, 2015). To both investigate and capture potential scale effects, here, we consider a collection of boundary definitions of differing spatial scales covering a range from administrative to statistical and synthetic boundaries.

Administrative and statistical. One of the simplest approaches to defining cities is to follow the administrative jurisdictions.[14] In the UK, the local authority units most often represent the administrative boundary at the closest appropriate scale for a definition of city. Referring back to the theoretical framework, administrative units, however, tend not to uphold the assumption of a homogeneously mixing population since their area may not directly correspond to a single independent city or market area due to the arbitrary nature of their definition, Figure 3.4. On the other hand, they satisfy the infrastructure network requirements of the framework, at least theoretically, and do represent governing bodies potentially capable of co-ordinating the required planning and infrastructural strategies. In this work, we consider the following administrative and statistical boundaries

[14] Understandably, these appear to be a favorite for policy and planning purposes in England and Wales where they are either used directly or as constituting units of larger blocs, e.g. Local Enterprise Partnerships (LEPs) or the devolved city-regions and combined authorities.

- local administrative units level 1 (LAU1)
- nomenclature des unités territoriales statistiques 3 (NUTS3)

These units, however, tend to become wholly inappropriate when considering London as they break down connected parts of the city to maintain statistical uniformity. As such for these boundaries, unless stated otherwise, we would consider as one the merged constituting units in accordance with the extent of Greater London Authority.

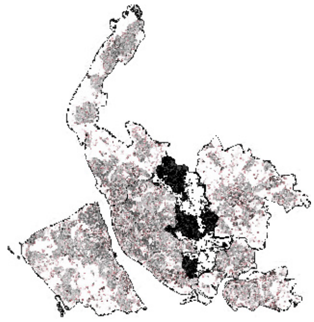


Figure 3.4: Schematic showing the Liverpool City Region and its constituent LAU1 units and their built footprint with the administrative unit of Knowsley highlighted.

Functional urban areas. An alternative to the statistical boundaries, which would address the potential incompatibility of the LAU1 units with the theoretical model's ideal mixing and mobility assumption within metropolitan areas, are city boundaries constructed thorough the consideration of commuting patterns and local labor market particularities at higher resolutions. Common to all approaches belonging to this family is a recursive algorithm in which smaller geographical units, e.g. output area (OA) in England and Wales, for which commuting data are known, are aggregated if a predefined portion of their commuter and resident population live and work within their combined territory. The aggregation is continued until all aggregated units meet an overall commuter/resident containment criteria (Coombes and Office for National Statistics, 2015). Consequently, these commuter-based aggregations should, by definition, meet the mixing population assumption inherent in the model unlike the arbitrarily drawn, in a functional metropolitan sense, administrative boundaries. It is worth mentioning that urban units on which Bettencourt's model is based, and from which the underlying body of the American empirical observations is derived, the metropolitan statistical area (MSA), belong to this category of boundary definitions to some extent as well, although not strictly constructed based on commuting patterns (US Census Bureau Demographic Internet Staff, 2018). In this work, we consider the following functional urban boundaries:

- Office for National Statistics travel-to-work area (TTWA)
- European Urban Audit functional urban areas (URBAUD)
- Organizaiton for Economic Co-operation and Development (OECD) Harmonized Functional Urban Areas

Although all three definitions use the same principal algorithm, there are few defining differences. The URBAUD and OECD definitions follow minimum population thresholds for delineating 'urban' areas resulting in a partial coverage of settlements (Eurostat, 2017; OECD, 2012), with those failing to meet the criteria excluded, limiting the number of available data points. Such an exclusion hampers an effective exploration of the existence and strength of the scaling patterns in a single urban network required prior to assessing the city performance within the SRM framework. This is while the TTWAs, which are England and Wales exclusive, provide full coverage of England and Wales land-mass. Moreover, to achieve the comparability within

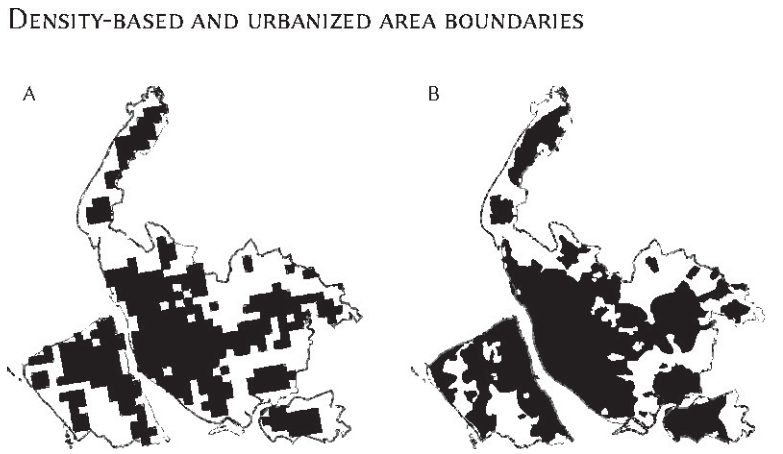


Figure 3.5: Schematic showing, A, the cell clusters with population density of 1400 $\frac{\text{prs}}{\text{km}^2}$ from a population grid of $1 \times 1\text{km}^2$ within the Liverpool City Region and, B, the contiguous urban cover clusters for the same region – note that when considering density-based boundaries, inset A contains multiple *city* units.

their dataset, OECD methodology utilizes a more lenient commuting threshold compared with that of the TTWAs (OECD, 2012, pp. 26-30) without localized considerations of the employment market applied to the TTWAs.[15]

[15] This is only of relevance when considering London metropolitan area. Unlike URBAUD and OECD definitions that leave London as one large functional area, in terms of TTWAs, areas of Heathrow and Slough are treated as a separate unit from the rest of London metropolitan region.

Density-based. Finally, the main criticism of the universality of the scaling exponents, especially those investigated in and for the urban system in England and Wales (Arcaute et al., 2015), involves the estimation of these exponents for boundaries extracted by clustering smaller units and their neighbors based on different population density thresholds. In this work, we also consider urban boundaries constructed through the aggregation of small neighboring units the population density of which exceeds a predefined cut-off value. The clustering itself involves another recursive algorithm in which a population grid, or alternatively other appropriately small statistical geographical units, e.g. ONS wards or OAs, is used, where cells are aggregated with their neighboring cells, 4 or 8 cells for a square grid (the von Neumann or Moore neighborhoods) and based on shared borders and/or certain distance thresholds for less structured geographical units, if the neighboring cells meet the minimum density criteria. It is perhaps worth mentioning that at certain density thresholds, around 1400 persons per square kilometers (Arcaute et al., 2015; Rozenfeld et al., 2011), clusters constructed in this fashion closely follow the contiguous urbanized area boundaries, Figure 3.5. To construct such units, we consider the GEOSTAT $1 \times 1\text{km}^2$ population grid which provides a pan-European grid with cell populations based on national census data (Eurostat, 2016). In this work, we consider the following boundaries for the following density thresholds:

- 100 persons per square kilometers (C100)
- 350 persons per square kilometers (C350)
- 500 persons per square kilometers (C500)
- 750 persons per square kilometers (C750)
- 1000 persons per square kilometers (C1000)
- 1400 persons per square kilometers (C1400)
- 3500 persons per square kilometers (C3500)

3.3.2. Data

As we have seen, three primary inputs, i.e. population, economic output, and urbanized area, for each city unit for a given boundary definition are required in order to examine the infrastructural needs of the urban system using the social reactor model.

Population. Although all the administrative and functional urban definitions used here are provided with regular population estimations, in order to minimize source and methodology variability and hence incomparability between boundary definitions, we use the GEOSTAT grid from the year 2011 as the base population layer used to estimate population for all units across all definitions. For density boundaries, these estimates are essentially the sum of population count in each individual cell attributed to the larger unit. For estimates in other boundary definitions, cell population is attributed based on the proportion of the grid cell area within the unit boundary of the other definitions with unit population summed as

$$N_j = \sum_i {N^i}_{cell} \frac{{A^i}_j}{{A^i}_{cell}} \tag{3.12}$$

where N_j is the population of j^{th} unit for a given boundary definition, ${N^i}_{cell}$ and ${A^i}_{cell}$ the total population and area of the i^{th} GEOSTAT cell, and ${A^i}_j$ the area of the grid cell i intersected by unit j.[16]

[16] Note that the method implicitly assumes a uniform population density over the area of each grid cell.

Economic output. Another reason to standardize population based on a single base layer is that, unlike population counts, estimates for economic output are not necessarily available for all boundary definitions. An important issue is then the availability of estimates for economic output, in the form of regional gross value-added and/or gross domestic product, at different scales and resolutions. While statistical bodies, such as the ONS, have detailed multivariate models to estimate wealth-related indicators, e.g. employment, income, etc., at higher resolutions using aggregate values measured at larger scale boundaries, e.g. LAU1s and NUTS3 regions, the utilization of these models tends to be cumbersome, time consuming, and would require a larger amount of information pertaining to the units. Alternatively, GIS-based methods using area- and population-weighted proportionalities provide a simpler and quicker estimation approaches.[17] These approaches involve a few simple steps. Firstly the layer containing the GVA estimates, specifically those calculated using the income approach that assigns values based on their geographical origin of production – aggregated and linked to their corresponding NUTS3 level areas,[18] is intersected with the population gird and each cell is assigned a portion of the GVA value according to

$$Y_{cell} = \sum_i \frac{Y_{NUTS3} \frac{N_{cell}}{A_{cell}} A_i}{N_{NUTS3}} \tag{3.13}$$

where the Y_{cell} is the total GVA assigned to a cell in the population grid, N_{cell} and A_{cell} the total population and area of the cell, Y_{NUTS3} and N_{NUTS3} the GVA and population of the intersecting NUTS3 areas respectively, and A_i the area of the portion of the cell intersected by the corresponding NUTS3 area. Subsequently, a similar procedure to that in Equation 3.12 is used to aggregate back the GVA values from the population grid to the desired city boundary layers discussed previously.

We should note, however, that the simplicity of this method might cause problems when aggregating back up to units slightly larger than the base population layer resulting in linear scalings or noise recordings (Smith, 2014), due to the nature of the simple population proportionality and the uniform density distribution assumption in Equation 3.13. While this should not cause significant problems within the scope of the work here, potential variations in output estimates are addressed in chapter B.

Urbanized area. While the total area, A, taken as that covered by the entirety of units' bounding box for each definition, can easily be estimated using any GIS package, we do in fact need the urbanized network area. CORINE land cover maps (European Environment Agency, 2000; NERC Environmental Information Data Centre, 2016) provide high-resolution and pan-European classification of land cover every dozen years or so. This enables an estimation of the urban network area, A_n, by considering the total area of the contiguous urban land cover[19] contained within each unit in different boundary definitions, Figure 3.6.[20]

[17] It should be noted that the OECD also utilizes these area-based methods for the estimations relating to their aggregation of urban areas (OECD, 2012, pp. 45-48).

[18] At the time, workplace-based estimates at NUTS3 boundaries were highest spatial resolution for output data. Although these remain the highest resolution units across Europe, the ONS has recently released experimental resident-based estimates at LAU1 boundaries.

[19] Contiguous urban cover refers to patches of primarily urban land cover that are less than a certain threshold apart, commonly 200-300 meters, and/or those natural covers that are surrounded by such patches (NERC Environmental Information Data Centre, 2016, pp. 26-39).

[20] In chapters 4 and 6, we would additionally use contiguous urbanized area (UA) as an alternative boundary definition similar to the density-based boundaries.

3.4. Concluding remarks

This chapter has provided a description, and in some instances a short critique, of the underlying theoretical model, tools, and datasets

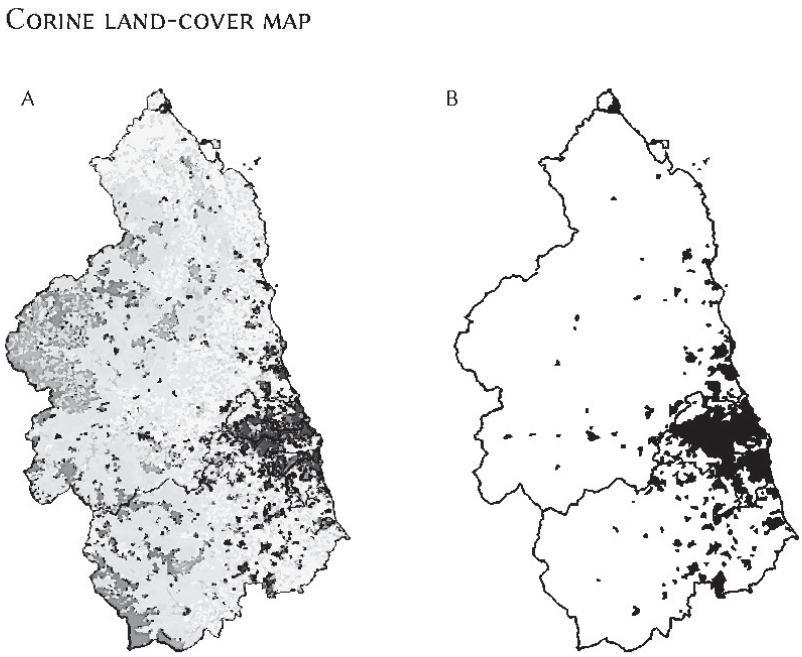


Figure 3.6: Maps showing, A, full CORINE land-cover map and, B, the corresponding contiguous urban area at a 200m threshold in the North-East Combined Authority.

that are used in this work to explore the broad objectives set out in chapter 1. Given the compartmentalized and self-contained nature of chapters 4, 5, and 6, in the current chapter we have simply focused on a succinct description of Bettencourt's SRM, which underpins the aforementioned chapters, and the procedures through which input data have been processed for these works. Methodological particularities for each chapter is hence addressed later on within the relevant chapter.

Finally, urban scaling frameworks, compared with their urban economic counterparts, provide a number of additional advantages. Most importantly, we reiterate that allometric frameworks are significantly more parsimonious and hence more practical in application. This enables power-law scaling models, such the social reactor model outlined above, unlike their New Economic Geography counterparts, to remain practical in circumstances where data is sparse and more agile when applied to an increasing number of cities and urban systems. Additionally, the few fundamental assumptions underpinning such models, as we have seen, are more general and avoid strong assumptions about individual behavior. As such, these models are not driven by individual behavioral assumptions and rather the empirically observable average-aggregate behavioral patterns of cities and the urban systems to which they belong. For these reasons, scaling models are computationally more tractable which allows further expansions without increased complexity.

4

A Multi-Scale Overview of Urban Performance

Material prepared for this chapter have been used in the following–

Arbabi, H., Mayfield, M., & Dabinett, G. (2019). Urban Performance at Different Boundaries in England and Wales through the Settlement Scaling Theory. *Reg Stud*, 53(6), 887-899.

4.1. Introduction

In this chapter we aim to offer some new insights on the effects of spatial scale on city performance balance and whether allometric arguments regarding connectivity and agglomeration remain valid across different scales and definitions of city boundary. As we have seen, social reactor model (SRM) provides an explicit formulation of the balance between economic output, the availability of mobility, and its associated costs incurred with reference to the actual physical extent of cities and the larger urban systems to which they belong. Such agglomeration frameworks, allometric or not, and much of their evidence are, however, based on Asian and North American urban systems. Consequently, England and Wales offer a particularly unique opportunity, in addition to our principal research questions set out in chapter 1, for the examination of these spatial effects within an urban system that is currently experiencing unique economic challenges (McCann, 2016, Chapter 3, pp. 121-124) with the ever-widening divide that exists between the productivity and economic output of the South of England and the other regions in England and Wales (Rowthorn, 2010).

The results, as we will see, signal a systemic lack of adequate mobility and accessibility for a large portion of city units considered at various spatial scales implying that lack of adequate mobility provisions is at the heart of a less-than-expected economic performance. While such effects are more easily noticeable at larger inter-city scales, the problem is reoccurring at smaller scales and intra-city boundaries. This suggests that although intra-city connectivity-based agglomeration strategies are fitting, when implemented alone they would only mask inadequate connectivity at smaller scales without addressing underlying causes of such under-performance. As such, transport infrastructure planning cannot simply be led by agglomeration theory principles being applied at a single spatial scale and more concurrent consideration of urban scales is needed. Crucially, the broader interpretations of the results point at a scale-induced hierarchy of urban connectivity that governs potential improvements needed at inter- and intra-city scales.

The rest of this chapter is structured as follows. In the next section, we first outline the relevance of English and Welsh urban system as a case study and it their transport infrastructure policy context. This is followed with a brief account of the methods and data used before establishing empirical agreement with the theoretical scaling model and examining the average city mobility and accessibility across spatial scales followed by their size-cost balance as compared with their idealized theoretical counterparts at these scales. In light of these comparisons, a brief discussion is provided focusing on the implications of inadequate mobility provision and the potential of connectivity-led agglomeration at different spatial scales.

4.2. England and Wales as a case study

Shaped and framed by the wider policy efforts stemming from the decentralization and devolution of certain powers to local entities in the form of combined authorities or city regions (Gardiner, Martin, Sunley, and Tyler, 2013), the infrastructure policy debate in England and Wales has been dominated in the recent years by the attempts to address an historic economic performance gap. For transport infrastructure at a national scale, these attempts have generally been envisaged as creating and enabling mid-sized cities to act as single economic units by providing inter-city transport infrastructures that reduce journey times encouraging agglomeration economies (National Infrastructure Commission, 2016a). The processes of devising these transport links, inspired by the inter-city rail connectivity in the Randstad and Rhine-Ruhr metropolitan region, has only been argued at a singular spatial scale (Transport for the North, 2016). Consequently, even though these policies have been studied in terms of their implications for infrastructure governance and funding (P. O'Brien and Pike, 2015), an explicit exploration of the scale effects on the size-cost balances and performance in England and Wales has largely been absent from both the policy papers and the larger academic debate.

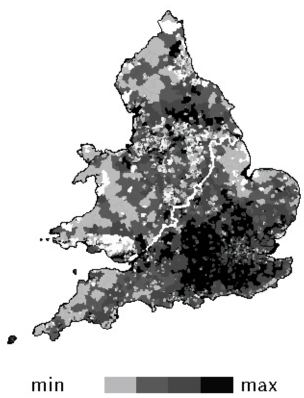


Figure 4.1: Map of England & Wales showing median household net weekly income in 2011 with Dorling's (2010) speculative dividing line overlaid.

4.2.1. UK regional economic divide

While part of a larger North-South division encompassing economic, political, cultural, and linguistic characteristics (Dorling, 2010), the regional economic performance of the UK, and more specifically that of England, has historically been beset by an economic gap dividing the country along a North-South border (McCann, 2016, ch. 1), see Figure 4.1. The economic geography of England and Wales, as a result, features a fractured spatial pattern. While the exact point in time when this divide in performance originated is contested (R. Martin and Gardiner, 2017) and suggested to stretch back as early as the mid 1800s (Geary and Stark, 2015), an as-of-yet still-widening division (Pidd, 2015; Stewart, 2015; Wang, 2016) between economic indicators of the northern and southern cities can be traced back to the late 1970s (Rowthorn, 2010).[1] As such, over the past few decades,

[1] Note that the Barlow Report (Barlow Commission, 1940) had forty years earlier noted an uneven pattern of economic and infrastructural development in the South-East to the detriment of the northern cities.

cities of the South, e.g. London, Slough, and Oxford, have enjoyed productivity levels greater than the national average while their five major counterparts in the North, i.e. Manchester, Leeds, Sheffield, Liverpool, and Newcastle, have consistently underperformed (R. Martin and Gardiner, 2017). This has culminated in such a disparity where the five northern cities, despite having in 2013 a combined population very close to that of London, had a combined gross value-added (GVA) less than half that of their southern comparator (Centre for Cities, 2015).

We should note here that despite the concerning impact of such regional imbalance, some theoretical and place non-specific frameworks, especially models originating from the New Economic Geography (NEG), might still see such disparities as a result of equilibrium conditions. From such perspectives, these conditions are seen as conducive to higher national growth and efficiencies (Gardiner et al., 2013) even though the empirical evidence one way or the other appears to remain inconclusive (P. Martin, 2005; R. Martin, 2008). Nevertheless, this regional economic divide and attempts at bridging it have in recent years come to dominate policy debate especially those relating to the hard infrastructure needed in the North to enable its cities to close the gap.

4.2.2. Infrastructure planning for spatial balance

The current infrastructural planning has come to include more solutions perceived to help make the North-South divide narrower. Most simplistically, these solutions, as far as hard infrastructures such as transport are concerned, have been framed through stylized agglomeration arguments. Taking the premise that bigger does more with less, the overall solution is seen as enabling the cities outside the South-East to form larger connected poly-centric metropolitan areas so that through their virtual combined population they form bigger and hence more productive *city-regions*. Although not an entirely new idea,[2] the current stream can be traced back to the Northern Powerhouse (NPH) program set out in a speech by the then Chancellor of the Exchequer, George Osborne (2014):

> Modern economists have spoken about the economic benefits when a critical mass of people, businesses and infrastructure are brought together in a large city. The whole is then greater than the sum of its parts. Our great northern cities represented here individually are quite small on the global stage – but combined they rival in size London or New York or Tokyo.

[2] A recent similar predecessor can be found in John Prescott's 'The Northern Way' (Parr, 2017).

> It was this opportunity to create a Northern Powerhouse that I identified earlier this year. I said that if we can bring our northern cities closer together – not physically, or in some artificial political construct – but by providing modern transport connections, supporting great science and our universities here, giving more power and control to civic government; then we can create a Northern Powerhouse with the size, the population, the political and economic clout, to be as strong as any global city.

In the proceeding years, this essentially inter-city connectivity driven agglomeration narrative has taken center stage as a crucial part of an infrastructural solution not exclusive to the northern cities and their Northern Powerhouse (HM Treasury, 2016) but also to those in the Midlands branded Midlands Engine (Department for Communities and Local Government, 2017). Although the 2017 Autumn Budget (HM Treasury, 2017) makes spending commitments towards intra-city mobility infrastructure as well, the overall solution is articulated as implementation and improvement of the inter-city transport infrastructures, almost exclusively in the form of reductions of the journey-time and improved and increases of service capacity and frequency of passenger rail, that would, as formulated by Osborne, enable these regions to function as a single economic unit.[3]

This process of regional aggregation, and/or perhaps forced economic agglomeration, is, however, taking place against a backdrop of lacking coherent planning strategies (Arcaute et al., 2016; Centre for Cities, 2015). In a comparison of the European Union nation states, Wang (2016) identifies England among a very small number of member states that do not develop a long-term 'strategic spatial plan' at a national macro-scale.[4] A recent review by the UK's own National Infrastructure Commission (2016b) has evaluated the national planning practice not only as lacking long-term strategy but also suffering from 'siloed decision-making' processes. This lack of planning strategy and a predilection for mega-projects is also noted publicly. Using rail services and the High Speed 2 as an example, Wolmar (Wolmar, 2016) points out how a rigorous and comprehensive assessment method for transport infrastructure planning is lacking, and consequently, small scale projects that would potentially address local needs more appropriately and multi-dimensionally are set aside in favor of grand business cases that rely on a singular increase of capacity and reduction of journey times or are alternatively based on mimicking supposedly successful approaches from other regions without a coherent strategy to assess and address wider needs.

This tension between large-scale and local projects highlights a

[3] Even without this drive to increase and facilitate inter-city connectivity, when considering the travel-to-work area (TTWA), regions of economically contained settlements, individual units have been gradually vanishing absorbed by their neighboring areas at a rate of three to four a year (Schifferes, 2015).

[4] This as Wang (2016) himself points out is extremely uncharacteristic of the otherwise highly centralized national government and unlike the other three nation states of the Kingdom that do in fact maintain such plans.

further disconnect in terms of the spatial scales used to examine, determine, and articulate both the infrastructural problems and solutions. Even though the macro-scale national North-South divide has been primarily framed as a connectivity problem and the solutions for it as meso-scale inter-city rail transport (Midlands Connect, 2017; Transport for the North, 2015), these meso-scale regional approaches are argued based on an arbitrary choice of spatial scale. This is often due the limitations posed by the availability of data or the relevance of the boundary for administrative purposes. As a consequence of a singular choice of scale, these meso-scale arguments are often blind to connectivity/productivity circumstances at alternative spatial scales and boundaries. Addressing this would require a comprehensive understanding of the interdependency of the urban form and its inhabitants, the infrastructure connecting them, and their economic output and energetic consumption across spatial scales.

4.3. Scaling and size-cost balance

As suggested in chapter 3, the scaling formulation of cities according to social reactor model (SRM) can be used to offer categorical comparisons of cities and urban regions both against an idealized realization of cities, i.e. $D = 2$ and $H = 1$, and also against any specific performance balance as observed in one particular city, say, London. In this section, we start by demonstrating the extent of the agreement between the underlying assumptions and resulting predictions from Bettencourt's model by estimating the scaling exponents for the gross value-added (GVA) and urbanized land area for the different city boundary delineations we outlined previously. We then estimate values of G for each city following

$$G_j = \frac{Y_j \times A_{nj}}{{N_j}^2} \tag{4.1}$$

which is a rearrangement of Equation 3.3 and where G_j is the human production estimated for city j and Y_j, A_{nj}, and N_j are the economic output, urbanized area and population of city j respectively. A comparison of these estimates for cities within each boundary definition with the optimal G^* calculated for their idealized fully accessible counterparts is then presented with an examination of the infrastructural needs of cities at different spatial scales.

Exact calculation of the optimal G^*, however, requires knowledge of values for the model's various internal parameters, e.g. transport costs. Nevertheless, without needing to fully estimate these, a system-wide average G^* can be obtained by substituting the scaling expressions of Y and A_n in Equation (4.1)

$$G^* = \frac{Y_0 N^{1+\frac{H}{D(D+H)}} \times A_{n0} N^{1-\frac{H}{D(D+H)}}}{N^2} = Y_0 A_{n0} \tag{4.2}$$

where Y_0 and A_{n0} are the system-wide prevalence of economic output and urbanized area respectively. Estimating an idealized optimal G^*, now, requires an idealized system as a point of reference. For this, we estimate idealized Y_0 and A_{n0} employing constant gradient ordinary least squares (OLS) fits on the linearized form of Equation (3.11) using Bettencourt's theoretical ideal scaling exponents of $\beta_Y = \frac{7}{6}$ and $\beta_{A_n} = \frac{5}{6}$ which correspond to our ideal values for D and H.

4.3.1. Urban performance in E&W

In obtaining the baseline prevalence and exponent of the scaling relations for urbanized area and GVA with population in each boundary definition, we use OLS estimators on the linear log transformation of Equation 3.11. The larger numbers of the excessively small units, especially in UA and C100 boundaries due to small isolated built-up areas and the smaller density cut-off, however, would skew the tail of the power-law and hence result in inappropriate linear fits. To obtain true estimates for the scaling exponents and the prevalence values within a scaling and agglomeration framework, it is then rather important to discard units with such low populations where effects of agglomeration, both increasing returns and economies of scales, vanish which often does correspond with the urban/rural designations. As such, arbitrary minimum population limits, such as 500,000 used by (OECD, 2012) and 50,000 used in Arcaute et al. (2015), are often used to distinguish urban and metro areas from those that are rural.

In this work, to seek such a population cut-off, we use the common observation of power-law rank-size distributions for the population of city units within the same urban system (Rozenfeld et al., 2011) by estimating the minimum population count above which a relatively stable power-law could be thought to apply. The python package 'powerlaw' (Alstott, Bullmore, and Plenz, 2014) provides an implementation of the statistical method used to estimate such minimum values described by Clauset et al. (2009). The method estimates the cut-off by seeking to find the point where the power-law best fit to the data resembles the probability distributions of the actual measured values the most. We use the method as a means to delineate urban from rural. As point of comparison, the Department for Environment, Food & Rural Affairs' urban-rural classification of 2011

LAU1 units in England (Bibby and Brindley, 2014) is at its core based on a population count approach.[5] The classification identifies units with over 50% *rural* population to be rural, with some considerations of dwelling type, services, and geography. The classification uses a population grid, similar to that we use in this work but finer in resolution, and deems areas (not LAU1 units) with over 10,000 inhabitants as an 'urban domain', subject to the services and density criteria. For the exception of C100, C350, and UA boundaries, the population cut-offs we apply to the units are considerably larger. Nevertheless, from a purely statistical perspective of urban-rural differences, these smaller units in C100, C350, and UA would appear to follow the same organizational patterns of those much larger than they are and hence be compatible for comparison in our adopted scaling framework. This in essence leaves us with units that are *urban* or follow the urban order of the wider distribution from a statistical perspective. Figure 4.2 shows the population complementary cumulative distribution functions (CCDF) for different boundaries marking the minimum population cut-offs and the best fit power-law.[6]

[5] The classification deems about 28% of all 326 units as rural.

[6] Note that in CCDFs for NUTS3 and LAU1 boundaries, the furthermost unit to the right is the synthesized unit of England and Wales that has not been used to estimate the cut-offs in these two boundaries.

These estimated minimum population cut-offs for each boundary definition and the OLS estimations for the units with populations above them are included in Table 4.1. It can be seen that while the overall regimes for β_{A_n} and β_Y are in broad agreement with the expectations developed and observed by Bettencourt (2013) and Bettencourt and Lobo (2016), the OLS estimates for the majority of the boundaries for both properties fall much closer to unity. This is especially pronounced in the decreasing trend of β_Y estimates at larger scales as population density cut-off decreases. Moreover, the outlier nature of the estimated β_Y for NUTS3 units can be attributed to the statistical nature of the boundary. Unit boundaries for this definition have a tendency to encompass an aggregation of administrative units such that the populous and economically active 'urban' bisected and cut out from the surrounding units in order to maintain population count uniformity across units for statistical purposes.

These deviations from prescribed idealized values of the scaling exponents have previously been noted with exponents estimated for the UK lying much closer to unity rather than the expected values of $\frac{5}{6}$ and $\frac{7}{6}$ for sub- and super-linear scaling, respectively (Arcaute et al., 2015). The larger matter of the comprehensiveness of these particular estimates is part of a broader ongoing debate that also includes issues around the appropriate methods of defining the boundaries of cities (Masucci et al., 2015). These, however, do not affect the study presented here since the derivation of the performance balance measure set out previously is independent from the estimated values of the exponents.[7] Within the framework of the social reactor model

[7] It should be noted that the numerical value of the theoretically optimal G^* is not independent of the exponents observed for economic output and urbanized area. It is the overall maximization of $Y - W$ that does not depend on specific values of the exponents.

Population CCDF and minimum population cut-off

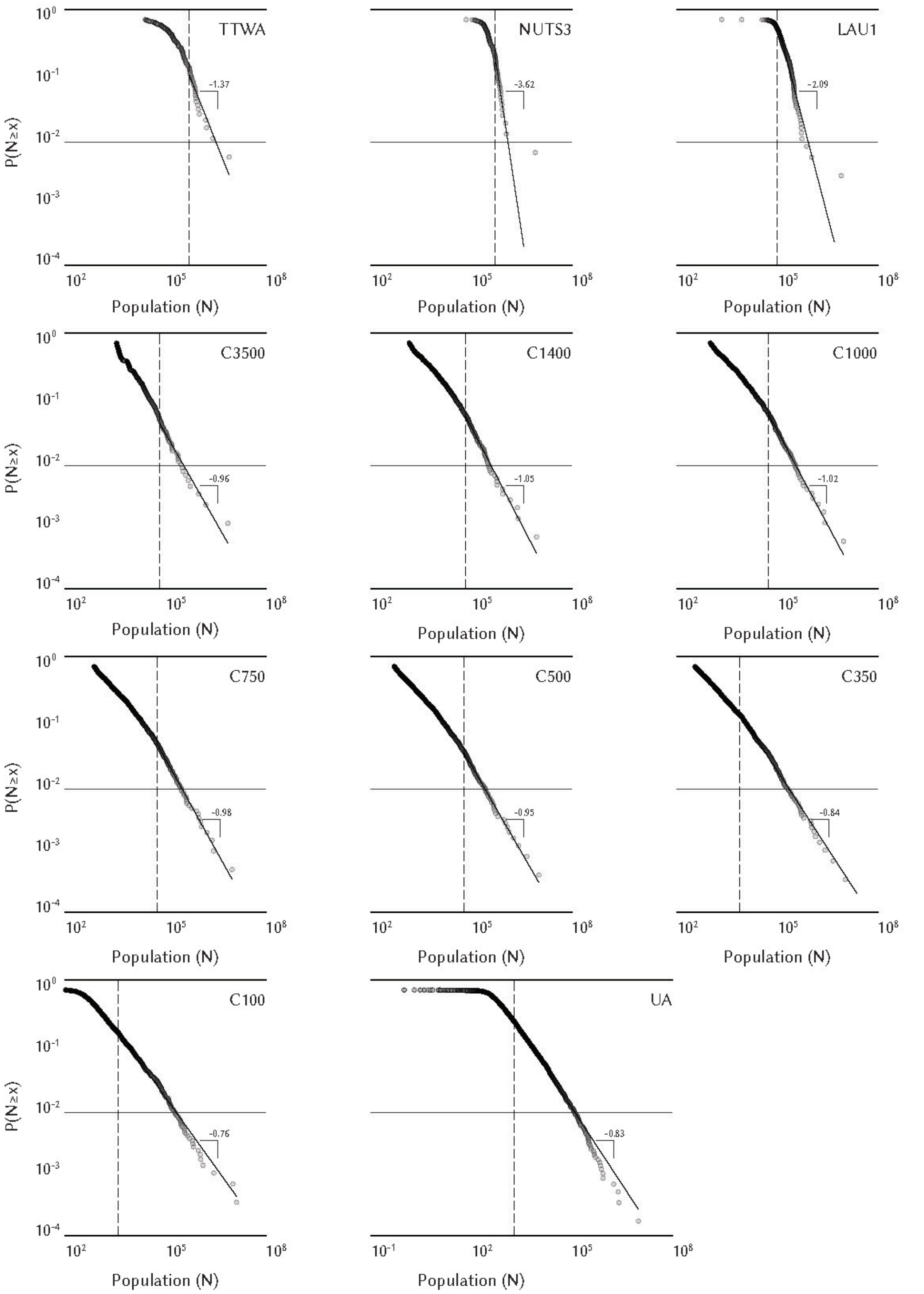


Figure 4.2: Population CCDFs and corresponding power-law fits – note that the dashed lines show the population cut-off values for the units above which a consistent power-law can be presumed to apply.

as discussed towards the end of the previous section, however, this prevalent linear scaling can be interpreted as a sign that cities in England and Wales on average exhibit a pattern of systematically and categorically impaired accessibility. As mentioned above, the extent of this lack of accessibility and mixing becomes increasingly larger at smaller population density cut-offs such as $100\frac{N}{\mathrm{km}^2}$ evident in the shrinking exponent estimates for the economic output. Nevertheless, the travel-to-work area (TTWA) boundary estimates for GVA and urbanized land area scaling exponents show a close match to those prescribed by the model, more or less appearing to uphold the mixing population assumption.

Table 4.1: Summary of the boundary definitions used and the estimated exponents for A_n and Y.

Boundary	**Units (N_{min})**[1]	β_{A_n} [95% CI]	β_Y [95% CI]	$R^2 \beta_{A_n}, \beta_Y$
C100	587 (3895)	0.94 [0.92, 0.95]	1.01 [1.00, 1.02]	0.97, 0.98
C350	481 (7627)	0.94 [0.93, 0.96]	1.02 [1.01, 1.03]	0.98, 0.97
C500	104 (59,698)	0.96 [0.94, 0.98]	1.02 [0.97, 1.06]	0.99, 0.96
C750	112 (57,698)	0.96 [0.94, 0.97]	1.02 [0.97, 1.06]	0.99, 0.95
C1000	120 (55,031)	0.95 [0.93, 0.97]	1.03 [0.98, 1.07]	0.99, 0.95
C1400	97 (67,495)	0.96 [0.93, 0.98]	1.03 [0.98, 1.09]	0.99, 0.94
C3500	49 (66,671)	0.95 [0.92, 0.99]	1.07 [1.00, 1.14]	0.98, 0.95
UA	1787 (1913)	0.94 [0.93, 0.95]	1.00 [1.00, 1.01]	0.91, 0.97
LAU1[2]	215 (101,355)	0.86 [0.82, 0.91]	1.02 [0.96, 1.08]	0.88, 0.84
NUTS3[2]	34 (499,766)	0.79 [0.70, 0.88]	1.29 [1.13, 1.45]	0.90, 0.89
TTWA	28 (510,149)	0.84 [0.76, 0.91]	1.14 [0.99, 1.29]	0.95, 0.91

1 Values in parentheses denote the minimum population cut-off for the smallest unit within each boundary definition when used for estimating the exponents.
2 Constituting boundary units for Greater London Authority have been aggregated and treated as one data point in these boundaries instead of treating boroughs or NUTS3 units as separate units.

Figure 4.3 shows the estimates of G for individual city units across the boundary definitions against population on logarithmic axes. It can be seen that despite the range of population size that is covered across the boundary definitions the estimates of G remain more or less independent of the population size (with an $R^2 \in [0.00, 0.06]$ suggesting $\frac{dG}{dN} \approx 0$) and within the same broad range across the different boundary definitions. This is in agreement with the assumption made by the model, with an overall median value of around $6.5{\times}10^6 \frac{£ \cdot \mathrm{m}^2}{N^2}$. The furthermost points to the right in each panel denote the different realizations of London and the Greater London Authority within each boundary definition. This regularity confirms the validity of the fourth assumption in Bettencourt's model formulated previously and the model's broader relevance in the context of city units in England and Wales.

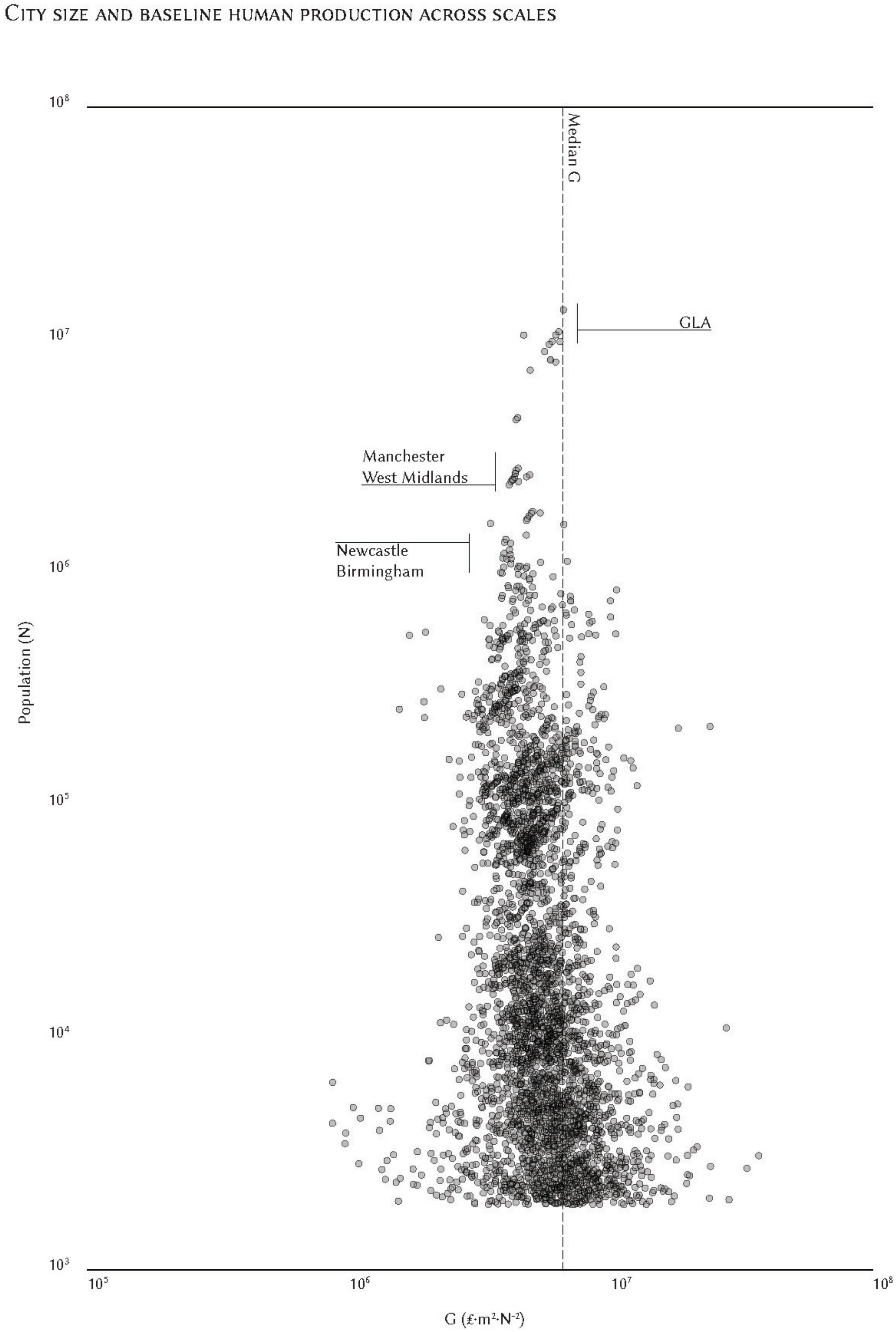


Figure 4.3: Superimposed plots of *G* and population for all boundary definitions with the overall median *G* highlighted.

4.4. Size-cost performance balance: mobility versus densification

Figure 4.4 summarizes the distribution of the ratio $\frac{G}{G^*}$, denoted as $\eta \equiv \log \frac{G}{G^*}$,[8] for city units in each boundary. From a first glance, it is clear that estimates of G do indeed tend to cluster close to the optimal value that maximizes the urban size-cost balance for idealized cities. A secondary observation can be made regarding the larger portion of the G estimates lying below the optimum highlighting a shortcoming in adequate levels of mobility and access in the city units across the boundary definitions used. This is more easily demonstrated by looking at the percentages of city units at different intervals of η where negative values indicate increasing lack of adequate mobility and mixing compared with the comparable idealized urban unit while positive values indicate higher needs for increased built-density, Figure 4.5. More than half of the units in density-based boundaries with cut-offs larger than $750 \frac{N}{\mathrm{km}^2}$, the two administrative boundaries, and the travel-to-work area (TTWA) show ratios below the optimum.

[8] Note that we use an arbitrary range rather than the absolute $\eta = 0$ when interpreting optimality allowing for minor variations about the empirically designated G^*.

Distribution of η across spatial scales

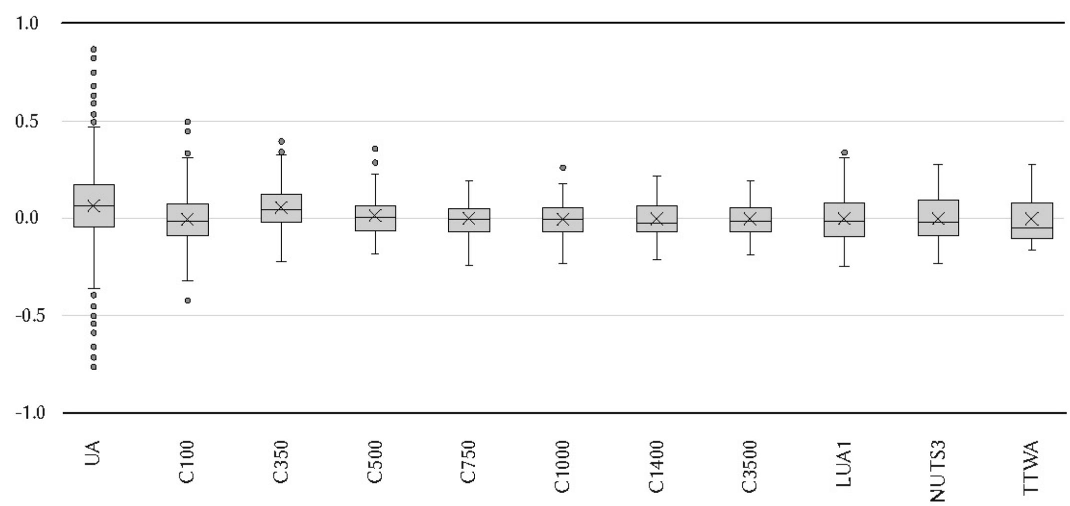


Figure 4.4: Box-chart showing the distribution of $\eta \equiv \log \frac{G}{G^*}$ within each boundary definition.

A cursory inspection of the units in the UA and C350 boundaries, which exhibit larger portions of cities with $\eta > 0$, indicates these larger portions consist of city units often of a small population that are near larger units or in close proximity of a number of other similarly small units where the economic output is effectively not a product of the interactions within single individual units and would involve interactions and commutes between units or to larger nearby conurbations. This can be verified by estimating η for city units dis-

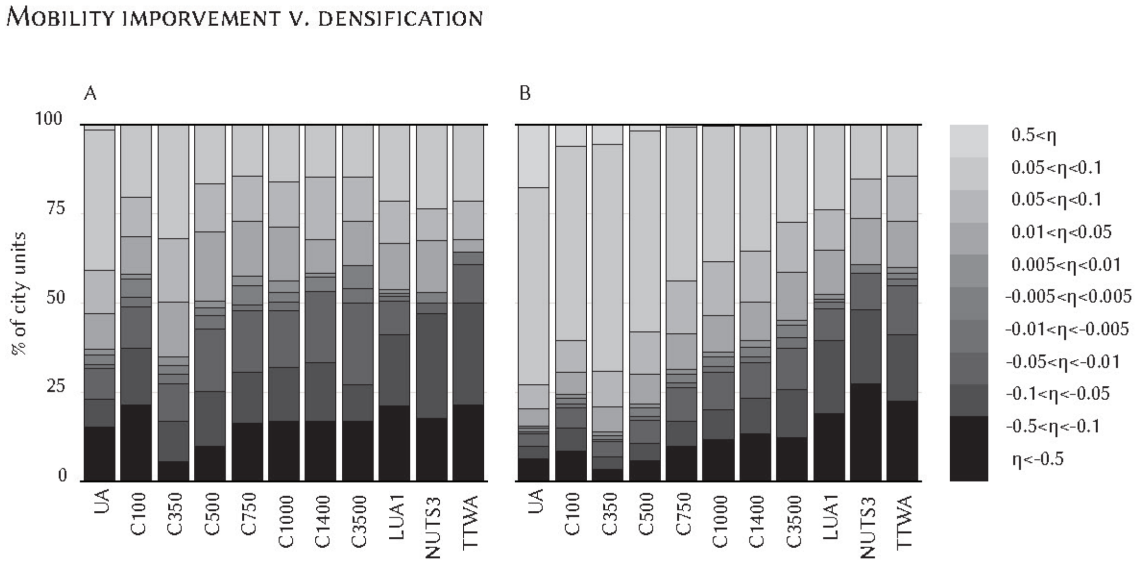


Figure 4.5: Bar charts showing the percentage of city units within the indicated range of η – left, city units above the population cut-offs in Table 4.1, right: all city units.

carded previously with populations below the minimum cut-offs indicated in 4.1. For this comparison, we do not re-estimate the theoretical point of optimum anew rather we use the theoretical optimum obtained for the larger *urban* units to quantify the notional performance balance of all city units compared with that of the average urban ideal, right panel Figure 4.5. This extension results in increases in the portion of units with larger than optimal η ratios especially in boundaries that would include large numbers of small city units on the periphery of larger ones, i.e. C100 and C350. The move from the smaller density cut-offs in C100 to those in C3500 in essence eliminates the satellite commuter suburbs where, as mentioned, gains in GVA are not achieved over their own urbanized area.

Finally, we geographically contextualize this optimality comparison by mapping each boundary definition and the corresponding ratio estimates. Figures 4.6 and 4.7 illustrate these for C100, C500, C1000, C1400, and the other non density-based boundaries. Note that the maps show estimated η for all city units within each boundary definition and not just those above population cut-offs indicated in 4.1. The first visual pattern to be immediately evident, especially in the density-based boundaries, is the change from below optimal ratios to those over the optimum crudely separating the South-East from the rest of England and Wales. Another notable observation is that the units corresponding to the Leeds and/or its greater city-region are the only major urban centers in the North exhibiting $G > G^*$ and as such, the only northern urban core indicating a need for densification to improve its size-cost balance rather than improvements to the intra-city transport similar to the rest. Additionally, subsequent disaggregation of the larger city unit of the North in the density-based boundaries as the density cut-off is continually raised from 100 to 3500$\frac{N}{\text{km}^2}$, does not appear to affect the identified size-cost balance where a need for better mobility persists despite the changing scales.[9] These remain largely stable even when comparing the corresponding boundaries in the non-density-based boundary definitions in Figure 4.7.

[9] Note that in C100 the combined areas of Liverpool through Leeds and then downwards through Nottingham are identified as a single city unit.

4.5. Chapter discussion

The planning policy in England and Wales is being driven with the emphasis on connecting the under-performing cities through improved transport infrastructure. As mentioned, this is seen as fundamental in enabling these regions to perform as a single functional economy and as such contributing towards the re-balancing of the national economy (National Infrastructure Commission, 2016a). These have precipitated in transport, specifically inter-city connectivity, building up the largest portion of the infrastructure pipeline 2017 onwards with project prioritization focused on reducing current travel time and reacting to the existing capacity demand while identifying the city regions with the highest economic opportunity associated with their inter-city connection (Infrastructure and Projects Authority, 2015).

The concluding observations from the previous section, however, noted a persistent lack of adequate mixing, or in other words a need for an improvement in the extent of the mobility provisions, in the majority of these regions regardless of the scale at which city regions are considered from LAU1s to larger TTWAs or density-based units. This is important when considering the generic recommendations borrowed from agglomeration theory regarding inter-city transport policy. The overall transport and connectivity focus of such insights appears in agreement and supported by the SRM's interpretation of the current size-cost balance in England and Wales across spatial scales. The inter-city focus of stylized agglomeration principles, however, ignores the overall performance balance, as formulated by Bettencourt (2013), and as such infrastructural needs across smaller scale boundaries similar to those depicted in Figures 4.6 and 4.7.

As an illustration, considering the C100 or C500 boundaries from Figure 4.6, center-to-center inter-city transport links connecting Liverpool, Manchester, Sheffield, and Leeds can be seen as beneficial. They would improve the performance balance as these regions appear as a single metro region with an apparent lack of appropriate

Figure 4.6: Maps of density-based boundaries color-coded based on the range of η. From left to right C100 and C500 at the top and C1000 and C1400 at the bottom – Contains National Statistics and OS data © Crown copyright and database right 2018.

Figure 4.7: Maps of density-based boundaries color-coded based on the range of η. From left to right UA and acrLAU1 at the top and NUTS3 and TTWA at the bottom – Contains National Statistics and OS data © Crown copyright and database right 2018.

connectivity plaguing their size-cost performance that an inter-city mobility scheme could to some extent remedy. All the while, the individual incarnations of cities building up these areas in the other boundaries, for the exception of Leeds, also show the same requirement for better mobility and transport across smaller areal extents. This is indicative of a lack of accessibility at different levels starting from within the high-density core areas, e.g. those in C1400, and persisting at larger scales, e.g. those in TTWA or C100.

With this in mind, transport-led agglomeration, as it is often articulated as facilitating connectivity between major city centers, involves mobilizing populations into city units that may individually not have the transport capability to provide for the efficient mixing and mobility that is implicit in agglomeration theory and conducive to the improved size-cost performance of the overall aggregated regions. Although such single-scale interventions could perhaps increase economic output nominally, the size-cost analysis suggests that they would do so to the detriment of the overall comparative balance at other scales. In England and Wales where the city centers have had the largest population growth in the last decade and accommodate the bulk of employment opportunities as the suburbs and rural areas provide the residential housing (Thomas, Serwicka, and Swinney, 2015), multi-scale, i.e. intra-city and inter-city, infrastructural interventions provided concurrently would seem more coherent. Considering practicalities such as a limited funding capacity, prioritizing policy interventions to start from smaller scales and moving on to larger ones would adjust the size-cost performance more effectively. This is so because improved mobility at an intra-city scale facilitates inter-city access while inter-city access would only increase demand on existing intra-city infrastructure.

We should, however, note that this is not to say that the inter-city infrastructure is not needed or to imply that it constitutes an entirely wrong strategy. In fact for any pair of cities where one exhibits $G > G^*$ and the other $G < G^*$, the SRM would project an estimate of G closer to G^* for the hypothetical and idealized city region which would have the sum of the pair's population, urbanized area, and economic output. Thus, assuming Leeds and Manchester areas comprise a well-connected metro region, the model would project a better-balanced size-cost performance for this hypothetical city, see the contrasting η estimates of the two cities in Figures 4.6 and 4.7. There is, however, an implicit assumption in the model used here that the resulting aggregated metro region is in itself a uniform urban conurbation providing for an ideal population mixing meaning for real-life examples balancing the performance of individual units would be required prior to a natural merging of the regions. In a sense, the agglomeration economies principle would perform as intended when the units are themselves performing well at smaller scales prior to connection at larger scales so that the aggregation helps to introduce the efficiencies and productivities of higher populations. These effects will not inexplicably overcome under-performance of the contributing cities if they are not previously addressed. This vital issue and importance of mobility at an intra-city scale is only often acknowledged in passing (National Infrastructure Commission, 2016a).

Finally, we might question the degree to which approaching G^* is desirable and practical. Despite its more tangible formulation of what essentially are *congestion costs*, Bettencourt's model aggregates all costs associated with population mobility over the infrastructure network. The energy dissipated and the overall size-cost balance, G, then have to include, for instance, fuel/energy source and type and cost bundled together. In a context where most mobility solutions are fossil-fuel intensive and concerns for the effects of climate change exist, maximizing the economic output for the transport energy lost becomes an imperative. In such cases, G^* embodies this maximization point and target. However, for the same targets, policy could focus on decoupling modes of mobility and transport from their fuel sources instead. As an extreme illustration, if similar levels of mobility could be provided through freely available public transport run by renewable energy sources the relevance of a $Y - W$ balance becomes diminished significantly. This would mean the optimum point of G^* may not eventually be a practical target especially in situations where cities are indicating estimates more than the optimum since escalating mobility costs will not have the same tangibly negative implications.

5

A Scaling Comparison of Three European Urban Systems

Material prepared for this chapter have been used in the following–

Arbabi, H., Mayfield, M., & McCann, P. (In Press). Productivity, Infrastructure, and Urban Density—an Allometric Comparison of Three European City-Regions across Scales. *J R Stat Soc A Stat.*

5.1. Introduction

We saw in the previous chapter that, on average, the urban network in England and Wales does in fact suffer from a lack of adequate provision of mobility and access in relation to the size-cost balance of its constituting city units more or less independent of the spatial scales at which these units are defined. These observations would to some extent reinforce and support the rationale for a transport- and mobility-orientated infrastructural approach when addressing the perceived regional economic under-performance in England and Wales.[1] In this chapter then, we aim to explore the universal transferability and applicability of such transport-driven measures from regions such as the Dutch Randstad and German Rhine-Ruhr metropolitan region to the English North.

[1] More broadly, chapter 4 supports a case for a multi-scale hierarchy of urban connectivity concerning potential improvements needed at inter- and intra-city scales.

As we will see, while inter-city connectivity arguments can be used when considering overall national performance of urban networks, particular inter-city connectivity solutions supported by stylized agglomeration-based arguments are not easily transferable from successful examples of poly-centric metropolitan regions in boosting under-performance of similarly sized regions elsewhere. Indeed, when considering size–cost balance, an examination of the needs from an urban scaling perspective can be made without requiring external comparisons. As such, continental case-studies, although very instructive, are not in themselves crucial in making a case for better connectivity in England and Wales. Such regional comparative approaches are, however, essential in identifying certain nuances which cannot be identified by looking at single-case data.

The rest of this chapter is structured as follows. The next section briefly contextualizes our expanded case study urban systems and providing a general comparative description of the three city regions and the broader urban systems to which they belong. This we follow with an outline of the variation on the main method from chapter 3 used here to enable the cross-country allometric comparison of the urban systems. We then present the results of this scaling comparison of the urban performance for Germany (DE), the Netherlands (NL), and England and Wales (EW) in the third section before proceeding with the comparison of the three city-regions and their constituting city units and urban zones. Finally, a brief discussion of the national and regional comparisons and their implications are presented in the last section.

5.2.
Continental inspirations

As previously stated, the Randstad and Rhine-Ruhr metropolitan region are often cited as typical examples of productive city regions with strong inter-city transport links. On the other hand, the North of England, as previously pointed out, is comprised of cities that are suffering significant economic under-performance despite their comparable urban size (Centre for Cities, 2015). This is symptomatic of a historic regional economic performance gap that appears to be unique to the UK (Dorling, 2010; McCann, 2016). The most recent attempts at addressing this economic under-performance in the northern regions of England has seen heavy reliance on stylized agglomeration-type arguments and the specific examples of the German and Dutch city regions to promote a larger city region, the so-called 'Northern Powerhouse', connected through inter-city passenger rail connections with decreased journey times and increased service frequency and capacity (National Infrastructure Commission, 2016a; Transport for the North, 2015). Although such arguments are inherently reliant on stylized agglomeration-type arguments, current transport schemes under consideration in England have particularly been influenced by and rely on examples drawn from the Randstad and Rhine–Ruhr metropolitan region. The case that has been made for such interventions by the relevant transport and infrastructure authorities draws specifically on the examples of the German and Dutch city regions when promoting a northern city region. Lacking from these arguments, however, has been a consideration of the compatibility and transferability of such connectivity-based scaling arguments between urban systems.

We first begin by providing descriptive comparison of the three countries and the city-regions of interest, drawing heavily on the data from Eurostat, Office for National Statistics, and the work by Swinney (2016). Table 5.1 shows a snapshot of population, employment, and output across the three city-regions. In spite of arguably similar populations of the Northern Powerhouse and Rhine-Ruhr metropolitan region, it is not difficult to see that, at least on the surface, the three city-regions can hardly be considered similar. In fact there is not a shortage of arguments pointing out that the two continental city-regions are not *similar* comparators for the North and would more resemble London (McCann, 2016, Chapter 2, pp. 50-73). What has prompted the use of these comparisons in policy documents and transport plans is perhaps the inspiration and example they would provide when attributing economic performance to a poly-centric connectivity. Given the geography of the North and from a comparable spatial scales perspective, looking at inherently poly-centric successful city-regions for potential connectivity lessons is arguably more intuitive than attempting the same with the connectivity infrastructure of London. It is useful to note, however, that since our overarching objective is to test the extent to which connectivity-based scaling arguments and insights are transferable between different urban systems and across spatial scales, we should in theory be able to use any urban systems for this comparison. Our specific choices of Germany and the Netherlands are thus convenient comparators that also provide for wider and thematically connected context.

Table 5.1: Comparative snapshot of the Northern Powerhouse, Randstad, and Rhine-Ruhr metropolitan region.

	Rhine-Ruhr Area	**Randstad**	**Northern Powerhouse**
Population	10.9 million	7.4 million	15.2 million
% of National	13.3	43.8	26.3
Number of Workers[1]	5.55 million	3.95 million	5.78 million
GVA[2]	310 billion	246 billion	297 billion
GVA per Worker	56,000	62,000	45,000
%Δ from National	8.7	5.3	-14.4

1 Values pertaining to employment and economic output are those recorded for the year 2013.
2 To account for different cost of living across the countries, figures are those of purchasing power standard (PPS) converted into British Pounds.

If the key differences underlying the higher productivities of the Randstad and Rhine-Ruhr metropolitan region were their inter-city transport enabling such agglomeration economies, then in an allometric framework including that of the social reactor model (SRM), we would expect distinct differences between the Northern Powerhouse and its continental comparators. These would exhibit themselves as differences of the system-wide scaling regimes, as evident from the national estimates for exponents β_Y and β_{A_n}, and also in the $Y - W$ balance, as evident in the distribution of η amongst the constituting city units of the city-regions. As we will see shortly, the overall English and Welsh urban networks do in fact exhibit a more pronounced systemic lack of adequate mobility when compared with their Dutch and German counterparts. Although our results support a case for better mobility and transport comparing the three urban networks regardless of the spatial scales, comparisons of specific city regions indicate a more nuanced interplay of productivity, connectivity, and urban density.

5.3. Normalized scaling

Before engaging in a one-to-one comparison of cities belonging to the urban network of England and Wales with their German and Dutch counterparts in terms of their size-cost optimality, we have to address the viability of such a comparison within the SRM's framework. From the previous chapters, it is clear that our determination of a system-wide representative point of optimality, G^*, is very much dependent on the baseline prevalences of economic output, Y_0, and urbanized area, A_{n0}. This means that a direct comparison of city units belonging to different urban systems requires the systems to have identical/comparable average characteristics. That is, however, not the case for the three countries of interest here especially when considering the scaling of economic output as the countries do not share a unit currency and as such unmodified scaling of Y across their cities would be expected to exhibit varying values of Y_0.

To enable a cross-country comparison, we follow Bettencourt and Lobo (2016) by normalizing economic output and urbanized area in each urban system. Here, this is done by normalizing city indicators by the idealized prevalence of the indicator in each system with this y-translation taking the form

$$\begin{cases} \ln Y_{j_T} = \ln Y_j - \ln {Y_0}^* = \beta_Y \ln N_j + \xi_{Y\,j} \\ \ln A_{n\,j_T} = \ln A_{n\,j} - \ln {A_{n0}}^* = \beta_{A_n} \ln N_j + \xi_{A_n\,j} \end{cases} \tag{5.1}$$

where Y_{j_T} and $A_{n\,j_T}$ are the normalized output and urbanized area for city i respectively, ${Y_0}^*$ and ${A_{n0}}^*$ the idealized fixed-gradient system-wide prevalence of output and urbanized area respectively, and $\xi_{Y\,j}$ and $\xi_{A_n\,j}$ the fluctuation terms from the strict scaling for city j. By means of this translation, the theoretical idealized model of output and urbanized area for each urban system now passes through the origin while leaving the scaling regime and exponents unchanged. As a result, the relative optimal baseline human production, G^*, for different urban networks is now similar and equal to unity. The normalization both enables a comparison of size-cost performance and a multi-system examination of the population scaling by investigating power-law fits to the combined data sample of the different urban networks.

5.4. Urban performance in Germany, the Netherlands, and England and Wales

Similar to chapter 4, we first begin with a brief overview of the boundary definitions examined across the three countries before examining the existence of power-law scaling and the empirical proximity of each country's urban network with Bettencourt's theoretical ideal. Figure 5.1 shows the population complementary cumulative distribution functions (CCDF) for different boundaries marking the minimum population cut-offs and the best fit power-law following the same methods as those implemented in the previous chapter.

From the outset, a glaring difference between the three countries concerns the minimum population cut-offs estimated for each of them at different boundary definitions. For the exception of C100 and C350 boundaries, the urban systems in Germany and the the Netherlands appear to follow a single coherent rank-size distribution over a larger portion of their smaller-sized units. This is in contrast with the urban networks constructed for England and Wales where a clear shift in the distribution exponent takes place over much larger population sizes. This to some extent suggests the existence of at least *two* English urban systems, with dynamics governing smaller settlements prevailing in much larger units when compared with German and Dutch distributions. Since we are interested in the productivity potential and size-cost balance of *urban* units we will continue to use and compare the tail to the right of the distributions when estimating system-wide responses.

Figure 5.2 and Table 5.2 show the OLS estimates for the GVA and urbanized area scaling exponents for each boundary and country.[2] As can be seen, the scaling of urbanized area and economic output do overall display a coherent sub- and super-linear relation with population respectively, regardless of the choice of country and/or urban network boundary definition. The extent of sub- and/or superlinearity of the relations, i.e. the strength of the economic productivity and infrastructural efficiency, however, does vary across countries and boundary definitions with Germany on average the most productive and efficient followed by the the Netherlands and England and Wales. By now, we know that from the perspective of the social reactor model, the deviations from the ideal exponents of $\beta_Y = \frac{7}{6}$ and $\beta_{A_n} = \frac{5}{6}$ towards unity indicate, on average, a system-wide lack of mobility, with $H < 1$, across all three countries with cities in England and Wales most affected. Nevertheless, the estimated scaling exponents, especially those of economic output closely trail the

[2] We reiterate that the use of simple OLS estimators is justified following the prior assumption and empirical observations that the scaling deviation term, ξ, follows a normal distribution centered on zero.

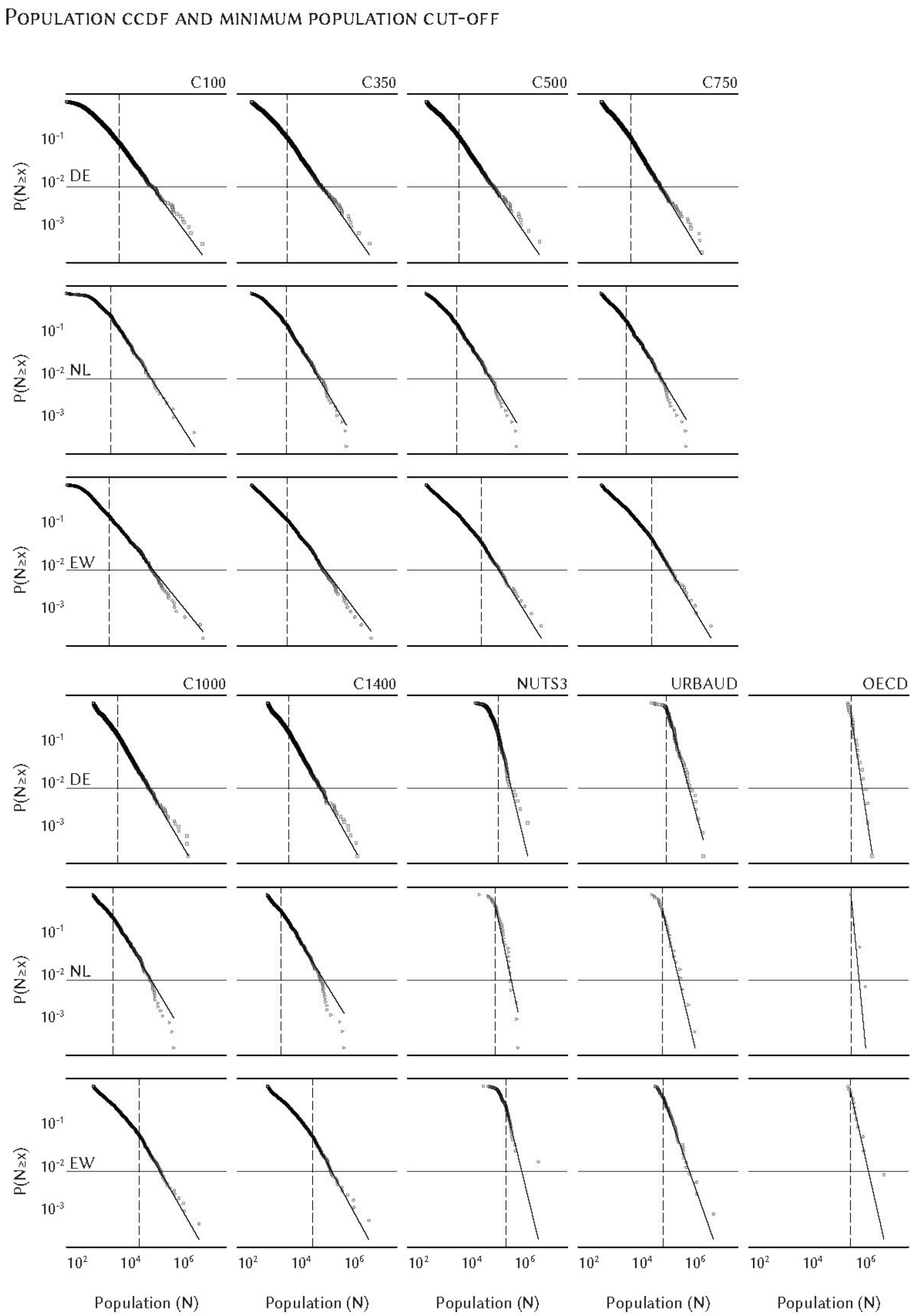


Figure 5.1: Population CCDFs showing population cut-offs and power-law fits.

theoretical ideal for the URBAUD and OECD functional urban areas which are the most directly compatible boundaries to those assumed within the model's assumptions (Bettencourt, 2013; Bettencourt and Lobo, 2016). Additionally, the complementarity of the output and urbanized area exponents for each boundary, i.e. $\beta_Y + \beta_{A_n} \approx 2$ implying $\frac{dG}{dN} \approx 0$, suggests that the model's third assumption also holds.[3]

[3] With R^2 of G against N averaging around 0.03 across different boundaries and countries.

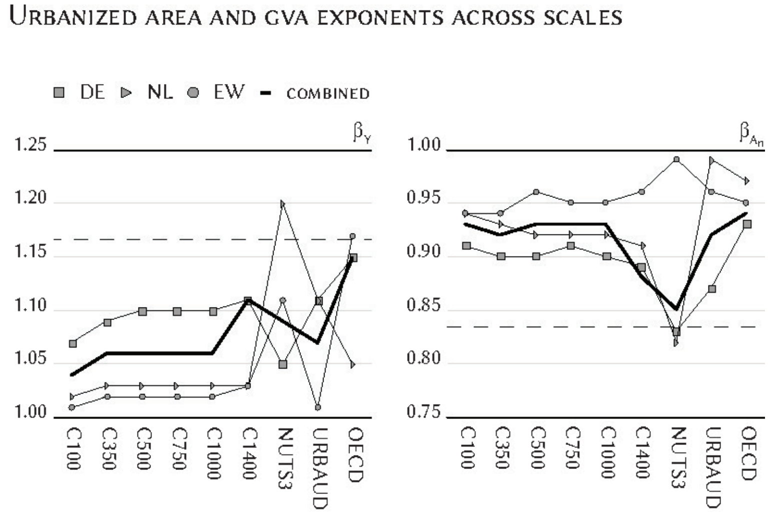


Figure 5.2: Plots showing the OLS estimated scaling exponents for each boundary, dashed line indicates theoretically ideal values for $D = 2$ and $H = 1$.

Similar to chapter 4, we see that from a comparative size-cost performance point of view, more than half of city units in England and Wales, regardless of the boundary, exhibit a need for better mobility to achieve their full economic potential, see Figure 5.3. The figure shows the percentage of units within a given comparative performance, $\eta(\equiv \ln \frac{G}{G^*})$, where again increasingly negative values indicate an increasing need for better within-unit mobility and transport while larger positive values an increasing need for built-area densification. It can be gleaned from the bar-charts that the size-cost performance appear more symmetrically distributed around the idealized optimum, $-0.02 \leq \eta \leq 0.02$,[4] when considering the aggregated distribution of performance balance for Germany and the Netherlands compared with those of England and Wales. When considering the boundary disaggregated estimates, the English and Welsh urban systems consistently exhibit a larger portion of units requiring better internal mobility and as such intra-urban transport solutions regardless of spatial scales, i.e. from core urban centers, e.g. C1400, to larger conurbations, e.g. C100 or URBAUD. This is while the boundary disaggregated picture across Germany and the Netherlands is more nuanced with Germany showing a slightly larger portion of units benefiting from densification efforts within C750 and C1000 in contrast with a larger need for mobility improvements for the same boundaries in the the Netherlands.

[4] Note that we use an arbitrary range rather than the absolute $\eta = 0$ when interpreting optimality allowing for minor variations about the empirically designated G^*.

Table 5.2: Summary of the OLS estimates for scaling exponents (rounded to 2 decimal place).

	Sample Size	β_Y	CI95%	R^2	β_{A_n}	CI95%	R^2
de							
C100	700	1.07	[1.05, 1.09]	0.95	0.91	[0.89, 0.93]	0.93
C350	965	1.09	[1.08, 1.11]	0.94	0.9	[0.87, 0.92]	0.86
C500	879	1.1	[1.08, 1.12]	0.94	0.9	[0.88, 0.93]	0.85
C750	768	1.1	[1.08, 1.12]	0.93	0.91	[0.88, 0.94]	0.86
C1000	827	1.1	[1.08, 1.12]	0.93	0.9	[0.88, 0.93]	0.85
C1400	17	1.11	[1.08, 1.13]	0.93	0.89	[0.86, 0.92]	0.84
NUTS3	402	1.05	[0.99, 1.10]	0.8	0.83	[0.77, 0.88]	0.70
URBAUD	94	1.11	[1.06, 1.16]	0.96	0.87	[0.80, 0.93]	0.88
OECD	24	1.15	[1.02, 1.28]	0.94	0.93	[0.80, 1.05]	0.91
Average		1.1			0.89		
nl							
C100	235	1.02	[0.99, 1.04]	0.96	0.94	[0.91, 0.98]	0.93
C350	246	1.03	[1.00, 1.05]	0.96	0.93	[0.89, 0.96]	0.92
C500	255	1.03	[1.00, 1.05]	0.96	0.92	[0.88, 0.95]	0.92
C750	272	1.03	[1.01, 1.06]	0.96	0.92	[0.88, 0.96]	0.88
C1000	296	1.03	[1.01, 1.05]	0.96	0.92	[0.88, 0.96]	0.89
C1400	339	1.03	[1.00, 1.05]	0.96	0.91	[0.88, 0.95]	0.9
NUTS3	40	1.2	[1.12, 1.28]	0.96	0.82	[0.71, 0.94]	0.84
URBAUD	34	1.11	[1.05, 1.16]	0.98	0.99	[0.91, 1.07]	0.95
OECD	5[1]	1.05	[0.82, 1.27]	0.98	0.97	[0.28, 1.66]	0.83
Average		1.06			0.92		
ew							
C100	587	1.01	[1.00, 1.02]	0.98	0.94	[0.92, 0.95]	0.97
C350	481	1.02	[1.01, 1.04]	0.97	0.94	[0.93, 0.96]	0.98
C500	104	1.02	[0.98, 1.06]	0.96	0.96	[0.94, 0.98]	0.99
C750	112	1.02	[0.98, 1.07]	0.95	0.95	[0.94, 0.97]	0.99
C1000	120	1.02	[0.98, 1.07]	0.94	0.95	[0.93, 0.97]	0.99
C1400	97	1.03	[0.98, 1.08]	0.94	0.96	[0.93, 0.98]	0.99
NUTS3	125	1.11	[1.05, 1.18]	0.90	0.99	[0.94, 1.05]	0.92
URBAUD	83	1.01	[0.95, 1.06]	0.95	0.96	[0.93, 0.99]	0.98
OECD	13	1.17	[1.05, 1.28]	0.98	0.95	[0.90, 1.01]	0.99
Average		1.05			0.96		
Combined Normalized							
C100	1145 (9501)[2]	1.04	[1.03, 1.06]	0.96	0.93	[0.91, 0.94]	0.94
C350	1668 (7847)	1.06	[1.05, 1.07]	0.92	0.92	[0.90, 0.93]	0.91
C500	263 (55,840)	1.06	[1.03, 1.10]	0.94	0.93	[0.90, 0.96]	0.94
C750	228 (65,987)	1.06	[1.02, 1.10]	0.93	0.93	[0.90, 0.96]	0.94
C1000	250 (60,383)	1.06	[1.02, 1.09]	0.92	0.93	[0.90, 0.96]	0.92
C1400	1022 (8577)	1.11	[1.09, 1.12]	0.95	0.88	[0.86, 0.90]	0.89
NUTS3	567	1.09	[1.06, 1.13]	0.87	0.85	[0.82, 0.89]	0.8
URBAUD	211	1.07	[1.04, 1.10]	0.96	0.92	[0.89, 0.95]	0.94
OECD	42	1.15	[1.08, 1.23]	0.96	0.94	[0.87, 1.01]	0.94
Average		1.08			0.91		

1 Beware the inappropriate sample size the effects of which are also reflected in the confidence intervals.

2 Values in parentheses denote the N_{min} used for the combined samples and obtained as set out in previous chapters.

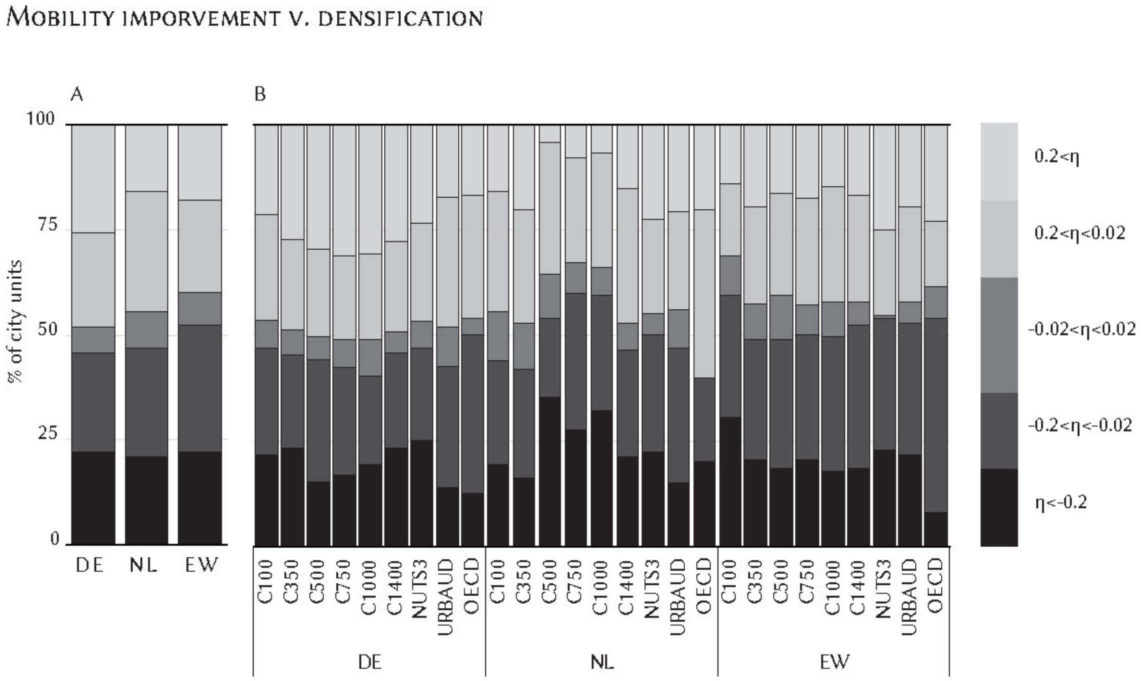


Figure 5.3: Bar charts displaying the percentage of city units in each country (A) and for each boundary definition (B) in the indicated range of η.

A combined interpretation of the comparative size-cost performance distribution and the overall scaling exponents estimated for each country suggests that all three countries are lacking in terms of urban mobility, albeit not to the same degree and not at the same spatial scales as previously noted. Meanwhile, England and Wales is further burdened with an additional prevalence of inadequate intra-urban access and mixing among its city units that appears unique among the three countries in its spatial persistence despite England and Wales's similar exponent estimates to those of the Netherlands.[5]

[5] For the sake of completeness, it is worth clarifying that this comparison is one of the comparative agglomerative productivities gaging the increased benefits associated with increased size and hence deliberately ignores the overall size of each nation's economy and their productivity as would be captured through the output prevalence Y_0 and the integral number and population of cities in each country.

5.5. The Randstad, Rhine-Ruhr metropolitan region, and Northern Powerhouse

The current infrastructure plans in England and Wales, as previously mentioned, focus heavily on the implementation of an inter-city passenger rail solution, combined with improving journey times and frequency explicitly borrowed from the Dutch Randstad and German Rhine-Ruhr metropolitan region, connecting and transforming a handful of the country's northern cities into a virtual city of a larger effective size, i.e. the Northern Powerhouse (Infrastructure and Projects Authority, 2015; Transport for the North, 2016). The results presented in the last section, in principle, regardless of the choice of city boundary definition and scale, support an infrastructure strategy concentrated on improving internal transport and mobility connections both simply based on England and Wales's isolated scaling and as a comparison relative to the performance of the German and Dutch urban networks. These national comparisons, however, would not necessarily justify the appropriation of an explicitly inter-city mobility solution from Rhine-Ruhr metropolitan region and Randstad for implementation in the English Northern Powerhouse. It is also crucial to note here that this examination of η masks individual economic productivity and infrastructure efficiency performance. Since η only considers the overall balance of $Y - W$, it entirely possible for cities to compensate for deviations from ideal scaling in one indicator, say Y, through complementary deviations in the other, i.e. A_n.[6] In such a way, considering Equation 5.1, a city unit with lower than ideally expected economic output, $\xi_{Y_j} < 0$ for $\beta_Y = \frac{7}{6}$, can compensate by incorporating a larger effective urbanized area, $\xi_{A_{nj}} > 0$ for $\beta_{A_n} = \frac{5}{6}$, in order to keep the overall G close to the point of optimal.[7] This leads to cities where despite a balanced cost-size performance economic under-performance may still be prevalent when compared with others.

[6] See for reference Figure 3.2 and the margin note on page 39.

[7] ξ, also known as a scale-adjusted metropolitan indicator, in essence captures deviations from the scale-invariant agglomeration behavior that are particular to individual cities (Lobo, Bettencourt, Strumsky, and West, 2013).

Consequently, we shift our focus to only those units within these three regions looking not only at their individual size-cost performance but also their deviations from an idealized expectation of output and urbanized area and that of the overall city regions they belong to by considering the hypothetical unit of their combined size summing their population, output, and urbanized area. We use two different approaches in defining the extent of the three regions and thus their constituting city units, one adopted from (Swinney, 2016) corresponding to an aggregation of NUTS3 administrative units and also representative of the *planned* Northern Powerhouse and the other based on the extent demarcated by the largest corresponding C100 units in each region, Figure 5.4. It is interesting to note that there is more agreement in the geography of the Randstad and Rhine-Ruhr metropolitan region defined either administratively or through urban proximity, i.e. single largest contiguous unit at a $100 \frac{prs}{\mathrm{km}^2}$ threshold, than there is between the two definitions of the Northern Powerhouse.

5.5.1. A regional comparison

Proceeding with our results, Figure 5.5 compares the size-cost performance of each region aggregated from units at each boundary definition and its overall deviation from the idealized output and urbanized area scaling. The dashed diagonal represents an optimal size-cost performance, $\eta = 0$, with the shaded areas corresponding to $-0.02 < \eta < 0.02$ and $-0.2 < \eta < 0.2$ similar to those in Figure 5.3. Note that due to the overall similarity of city units and scaling regimes for the density-based boundaries, from this point forward, nation- or region-wide aggregation of all units refers to all units within C100, C500, C1000, C1400, NUTS3, URBAUD, and OECD excluding the remaining density-based boundaries. Although this was done to minimize the double counting of city units the boundary of which does not change greatly from boundary to boundary while maintaining representation of scale changes, the exclusion does not significantly affect city distributions and results presented in Figure 5.5. Comparing only the size-cost performance of the regions, not much difference could be discerned between the Randstad, Rhine-Ruhr metropolitan region, and Northern Powerhouse. The majority of their different realizations indicate a need for better internal mobility and mixing, which would include inter-city mobility between their constituting city units, regardless of the choice of boundary definition or their overall extent despite the existing inter-city passenger rail infrastructure in the Randstad and Rhine-Ruhr metropolitan region. Out of the three, however, the Randstad shows a larger qualitative variation in estimates depending on the choice of boundary definition with realizations summed exclusively over constituting NUTS3 and URBAUD units indicating a need for densification. A similar need can only be seen for a Northern Powerhouse comprised from the OECD units within the C100 regional extent with no rendition of Rhine-Ruhr metropolitan region exhibiting $\eta \geq 0$. Meanwhile, the comparison would suggest that size-cost performance is already relatively optimal for both the Randstad and the planned Northern Powerhouse when aggregating OECD units despite glaring differences in the mix of cities involved in the two variations of the Northern Powerhouse. The consideration of the scaling deviations, on the other hand, highlights a pattern whereby the economic over-performance is correlated with denser built-areas.[8] From this perspective, despite seemingly larger imbalances of size-cost performance and a more pronounced need for better internal mobility the German and Dutch city regions outperform the Northern Powerhouse economically suggesting that policy measures differentiating them and the Northern Powerhouse and to be borrowed from the two are perhaps not simply those concerning inter-city mobility.

[8] $R = -0.8$ for aggregated regions and $R = -0.4$ when considering their constituting member units.

5.5.2. A sub-regional portrait of national differences

To complement the comparison of the three city regions and their home countries, we further calculate the percentage of cities within

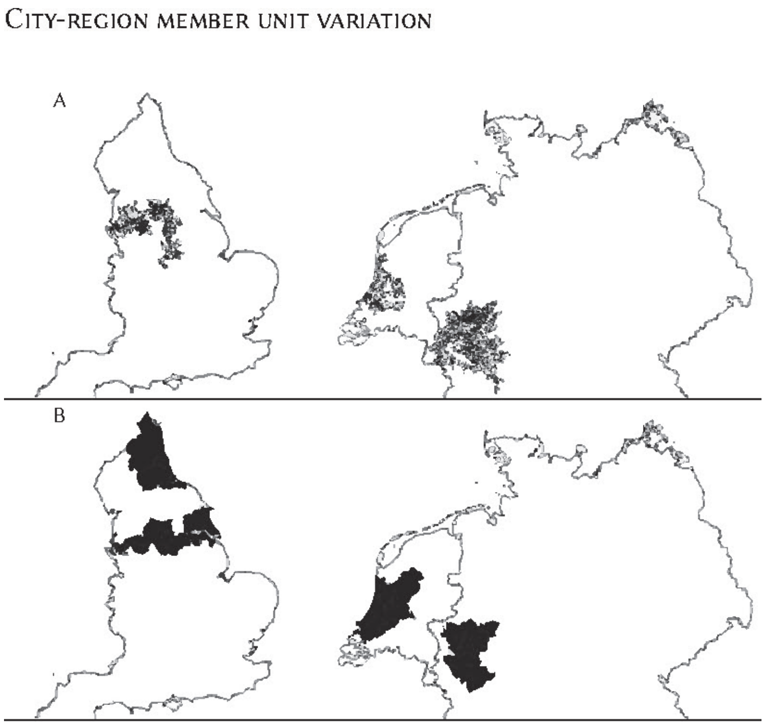


Figure 5.4: Maps showing the areal extents used for allocating units to the city regions – contiguous C100 units (A) and NUTS3 units (B).

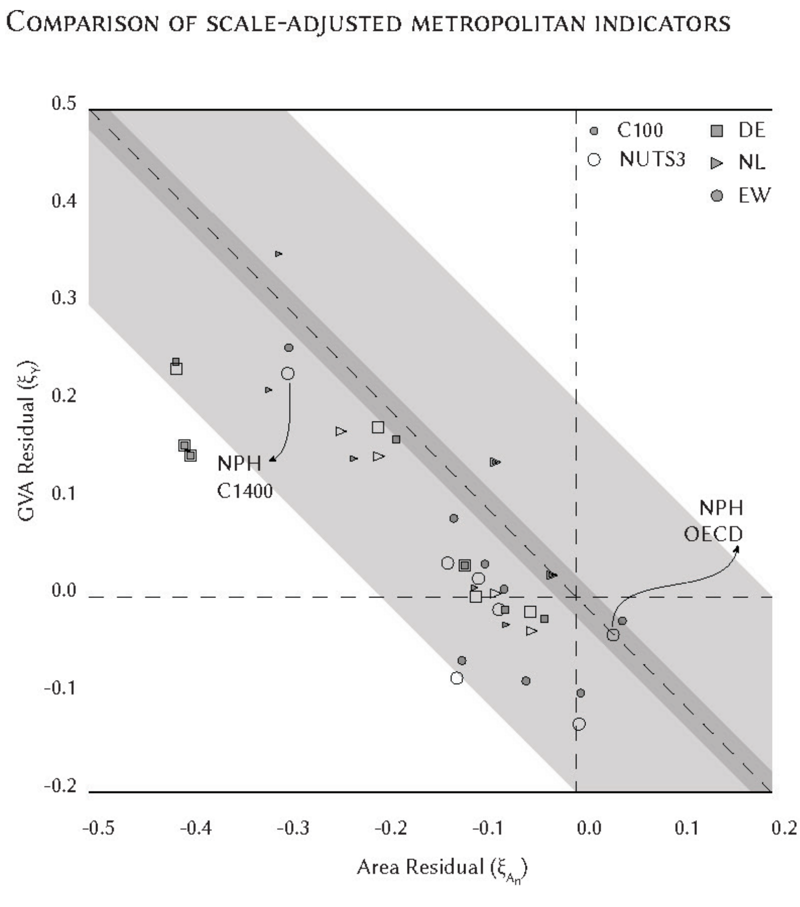


Figure 5.5: Scatter plot of output residual against urbanized area residual for Randstad, Rhine-Ruhr metropolitan region, and Northern Powerhouse assembled from units at different boundary definitions.

different ranges of ξ_Y and ξ_{A_n} building nation-wide and region-wide city distributions. Figure 5.6 shows discrete heat-maps with residuals for urbanized area on the x-axis and that of economic output on the y-axis and the cell color correlated with the percentage of city units lying within the cell. Note that the diagonal remains indicative of near-optimal size-cost performance.[9] The most noticeable difference between the nation-wide distribution of city units in DE, NL, and EW is the relative symmetry of the distribution about the diagonal in Germany and the Netherlands mirroring their distributions in Figure 5.3 with distribution peaks along the diagonal. Additionally, it is clear that these peaks in Germany and the the Netherlands are firstly units that are sparse and economically under-performing, i.e. those in the bottom-right quadrant, followed by those that are dense and economically over-performing, top-left quadrant. This is in contrast with the England and Wales national distribution where more than half of all units are within the lower triangle below the diagonal with the distribution peak pointing to cities that are economically under-performing despite their perceived density, bottom-left quadrant, with a size-cost balance in significant need of better internal mobility.[10] Of more interest is the difference between national and regional distributions. While comparing the composition of the Randstad and Rhine-Ruhr metropolitan region regions with the overall German and Dutch distributions highlights a shift of the distribution peaks from sparse economically under-performing city units to denser and over-performing ones, especially in the Randstad, where as a comparison of the Northern Powerhouse against the England and Wales composition reveals a slight increase in the portion of units that are both dense and under-performing.

[9] Note that although the diagonal from Figure 5.5 does pass through the diagonal of the matrices, the cells might also contain many units much further away from $\eta \approx 0$.

[10] It is worth pointing out that the top-left quadrant for the England and Wales heatmap is to some extent occupied by the city units in the South-East.

5.6. Chapter discussion

Similar to the results, we begin our brief discussion and round-up of the results with the national comparison. Model interpretations of the comparison of the scaling regimes governing the economic output and urbanized area in the urban networks of the three countries point to system-wide lack of adequate internal mobility and accessibility as fundamental to the lower productivity elasticities of the English and Welsh urban system compared with that of Germany. As such, while the findings from the comparison between the three countries' urban networks are consistent with expectations, the the Netherlands and Germany as national comparisons are not necessary in arguing for better transport in England and Wales. In this manner, simply assessing England and Wales's urban network in isolation, as already undertaken in chapter 4, with respect to the

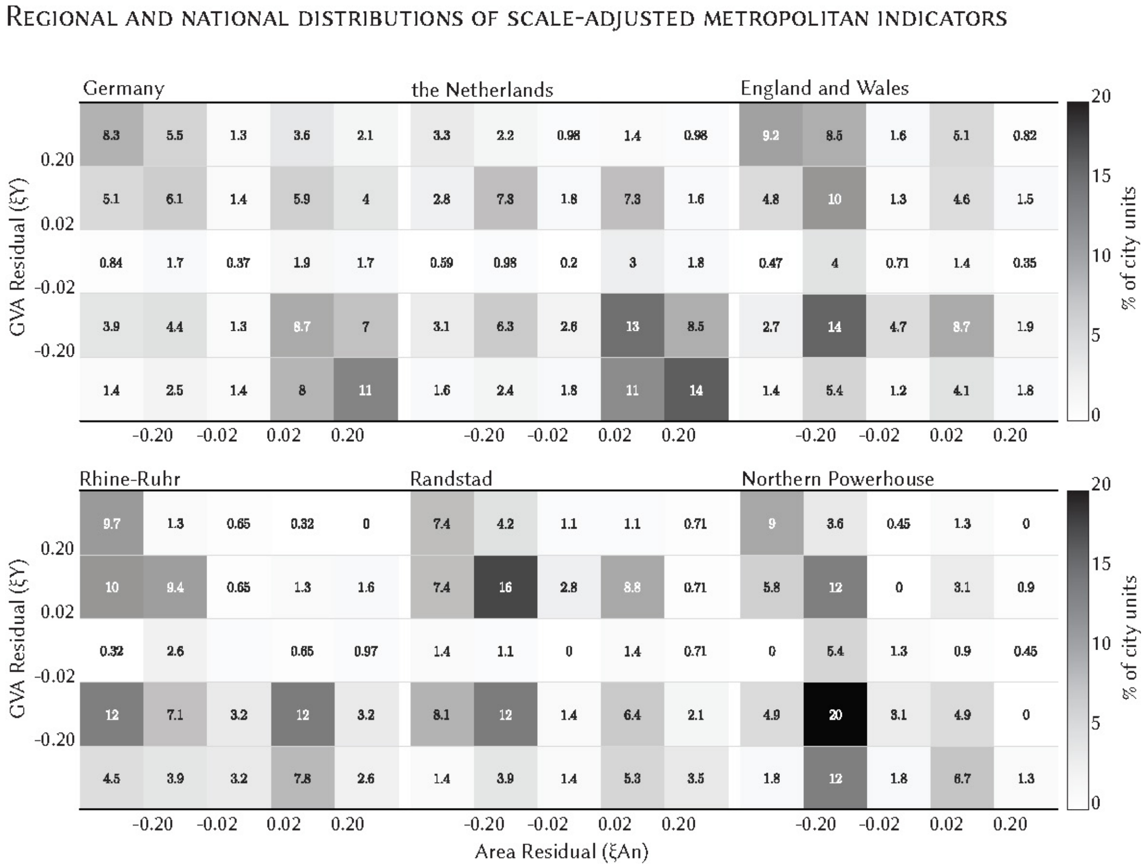


Figure 5.6: Heat-maps showing percentage of city units across all boundaries for each residual cell.

SRM's ideal could have supported a case for the deployment of better transport and mobility infrastructure, albeit those mostly of an intra-city nature, for boosting national economy and by extension that of the northern cities from an agglomeration point of view.

Moreover, a comparison of the scaling exponents estimated at the URBAUD boundary definition shows both German and Dutch urban networks exhibiting increasing returns to scale for economic output in contrast to the near-linear scaling regime in England and Wales. This is in spite of a similarly linear scaling of urbanized area observed for both the Dutch and English urban systems. It could consequently be argued that, in addition to the connectivity and mobility factors influencing the development and growth of the urbanized area and output productivity, a wider range of policy differences should be taken into account when explaining the disparity between the economic productivity of the three countries. In other words, although one might be able to extract transferable policy drivers from comparisons with better performing urban networks such as those of Germany and the the Netherlands, a singularly inter-city transport driven argument would not be the root solution/driver at which to arrive. The regional examination of the Rhine-Ruhr metropolitan region, Randstad, and Northern Powerhouse further reinforces this.

We have seen that on average the Randstad and Rhine-Ruhr metropolitan region are comprised of individual units that themselves economically out-perform the theoretically ideal expectation of their size regardless of the spatial scales. This is in contrast with the individual units building up either realizations of their English counterpart. The consideration of the aggregated regions with respect to the scaling residuals appears to suggest this to be more closely associated with the higher densities of the continental examples demonstrated by the comparison of the three regions at different boundary definitions, Figure 5.5, where the aggregated Northern Powerhouse shows considerably lower densities and by extension productivities. It is therefore notable that the only comparable over-performance of a Northern Powerhouse unit occurs at C1400 boundary definition, which is also its only realization of a comparably dense nature. The same density-productivity trend is also seen for the comprising units of the Randstad and Rhine-Ruhr metropolitan region with a majority of units denser and over-performing in contrast to their national distributions. Meanwhile, the composition of the Northern Powerhouse is very much representative of the England and Wales in general.[11] This re-frames the under-performance of the northern English units not as a regional problem but one at a national level. Nevertheless, the aggregate regional comparison, in contrast to the current transport-led infrastructural program, would suggest a need for further densification in Northern Powerhouse using the same agglomeration-based principles.[12] On a related note, we pointed in passing to the difference that exists between the geographic coverage of the *planned* Northern Powerhouse and its contiguously populated boundary, Figure 5.4. Although, insights from Figure 5.5 suggest that this territorial difference does not influence size-cost optimality significantly, such geographic proximity issues are bound to become influential when considering the practicality of implementing multi-scale mobility improvements and/or densification measures.

[11] It is interesting to point out that if we were more interested in London rather than the Northern Powerhouse, we would have seen a pattern similar to the Randstad and Rhine-Ruhr metropolitan region implying that perhaps there is as much to borrow from the density and transport profile of London for the Northern Powerhouse as there is in continental comparisons.

[12] A parallel case for urban density can also be made from a purely energetic perspective through examination of the broader trends of energy consumption versus urban population and density emphasizing potential energetic efficiencies associated with high-density urban living in England and Wales, see Appendix A.

Finally, an additional source of nuance is the implication of singularly deploying either inter-city mobility infrastructure or densification policies on the size-cost balance of the aggregated region, especially when factoring in the spatial scales over which the infrastructure is to be incorporated. Whereas the economic residuals appear to grow with multi-scale densification, i.e. a shrinking ξ_{A_n}, whether or not the overall cost-size performance remains near-optimal requires a balance between the two strategies to be reached. In this vein, Rhine-Ruhr metropolitan region can achieve higher potentials and size-cost balance through further improvements of mobility. The same is true for the Northern Powerhouse and Randstad across a majority of spatial scales. Under the agglomeration economies paradigm, therefore, improvements and extensions of the inter- and intra-city transport infrastructure become crucial not as the principle solution but as the complementary measures needed to maintain appropriate levels of mobility and hence size-cost balance as any of the regions densify as a whole, across all or a given boundary definition, towards the top-left quadrant in Figure 5.5.

6

Development Logic of City-Regions

Material prepared for this chapter have been used in the following–

Arbabi, H., Mayfield, M., & McCann, P. (In Press). On the Development Logic of City-Regions: Inter- Versus Intra-City Mobility in England and Wales. *Spatial Economic Analysis*.

6.1. Introduction

As we have repeatedly mentioned in this work, agglomeration-based arguments are used to support a case in favor of inter-city transport infrastructure and connectivity. These broader attempts at bridging the economic performance gap that exists between the northern regions and London have framed this divide as a mobility problem (Osborne, 2014). This has resulted in use of similar stylized agglomeration arguments in favor of implementation and upgrades of the passenger rail infrastructure to increase capacity and reduce journey times. These transport interventions and region building efforts are envisaged to enable northern regions to act as a single economic unit leveraging their virtual collective size for higher productivities (Transport for the North, 2015).

Over the course of the last two chapters, we have, through the lens of the social reactor model (SRM), explored the size-cost balance of city units in England and Wales and their performance in comparison with their continental counterparts in Germany and the the Netherlands. We have seen that mobility does indeed appear to be at least one of the crucial factors in the North-South performance divide, although not necessarily at the same spatial scales and boundaries over which current large-scale infrastructure may be focused. Having observed the impact the choice of city units might have when considering the aggregated performance of the city-region they collectively represent in chapter 5, we now aim to investigate the effects of spatial scales and distance on the geographic patterns of transport-led agglomeration strategies from a cost-size perspective. Continuing to use Bettencourt's framework which provides an explicit formulation to identify key infrastructure interventions needed, i.e. densification or better mobility measures, to balance city performance, we expand on it by adapting a pseudo-hierarchical linkage clustering algorithm to pair city units with complementary infrastructural requirements where pairings mirror provision of inter-city transport links. This additionally allows us to investigate the robustness of such groupings by performing a co-occurrence frequency analysis examining the recurrence of specific city-pairs over different aggregation scenarios. As we will see, our findings here reaffirm our observations and interpretations from chapter 4 that there appears to be a persistent inadequacy of population mixing and mobility across intra- and inter-city scales, which are predominantly frequent and potentially better addressed over short or intra-city distances.

The rest of this chapter is structured as follows. In the next section, as usual, we initially provide a concise account of the pseudo-hierarchical clustering method and aggregation scenarios used before exploring the outcome of the various aggregation scenarios and the resulting city-regions. This is followed by a brief discussion.

6.2.
A pseudo-hierarchical clustering

The argument for city-regions connected through effective center-to-center transport is often put forward through agglomeration principles whereby higher productivities are expected to result from the increase in the effective urban size via the upgraded transport. From the perspective of the Bettencourt's model, however, such inter-city mobility measures would not exhibit their full potential when all the cities to be connected have already inadequate levels of mobility, $G < G^*$, internally.[1]

[1] This in essence is simply stating that delivering people more efficiently into cities that have internal mobility shortcomings is in itself inefficient.

An overall complementarity can then be seen between cities that fall on either side of an idealized point of optimum, G^*. Figure 6.1 provides an instance. Suppose that city-unit A, according to the social reactor model, requires further densification to address its size-cost balance relative to an assumed point of optimality, $\eta_A (\equiv \log(\frac{G}{G^*})) > 0$, and that its neighboring urban area, city-unit B, is suffering from a lack of internal mobility, $\eta_B < 0$. If we were to consider the performance of this pair as a single hypothetical unit $A + B$, which implicitly assumes provisions of mobility between the pair, then the resulting city pair would theoretically lie somewhere closer to the point of optimum, $\eta_B < \eta_{(A+B)} < \eta_A$, with on average a reduced perceived need for further infrastructural intervention as a result of an adjustment in units boundary.

Pairing cities to mitigate inftrastructural needs

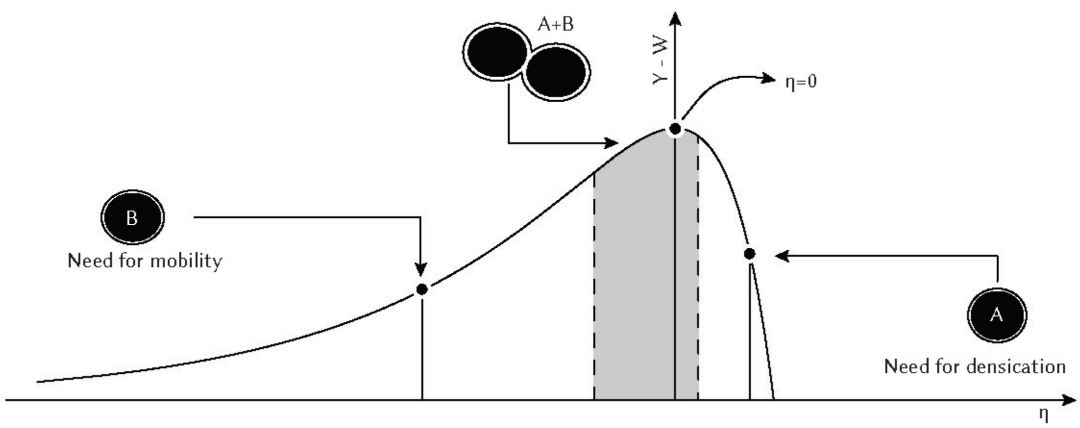


Figure 6.1: Schematic showing the individual comparative cost-size performance, $\eta(= \log(G/G^*))$, and potential combined city-pair performance in the shaded slice.

Consideration of combined cities could thus be thought of as a scale change in the local city boundary. This rearrangement of the boundary results in consideration of a city unit that has the aggregate sum of 'infrastructural extent' and 'economic output' of the parent units. For this hypothetical unit to then deliver on this aggregate infrastructural and output potential, i.e. in essence relocating closer to the stationary point on the $Y - W$ curve, would then require the aggregate inhabitants to have been provided with mobility levels, H, that is at least similar to the parent units across the combined area of the two. This is to say mobility levels, which are at least comparable to those connecting parent units internally, already exist or are subsequently provided across the two.

This can be used to systematically identify regional clusters where such agglomerative inter-city mobility upgrades provide a perceived closer-to-optimal size-cost balance. We employ a pseudo agglomerative hierarchical linkage clustering method (Murtagh and Contreras, 2012) grouping units together at each step where a distance function is expressed as

$$D(A,B) = |\eta_{(A+B)}| = |\ln \frac{G_{(A+B)}}{G^*}| \tag{6.1}$$

with A and B filling in for any set of cities or city-regions. The combined baseline human production, G_{A+B}, can be estimated through a rearrangement of Equation 4.1

$$G_{(A+B)} = \frac{(Y_A + Y_B)(A_{nA} + A_{nB})}{(N_A + N_B)^2}. \tag{6.2}$$

6.2.1. *Clustering scenarios*

We conduct our analysis over the England and Wales urban network for the same boundary definitions used originally in chapter 4. Table 6.1 shows summary description of the boundaries used. Clustering city units of each boundary in Table 6.1 according to the formulation above, however, would not account for the geography of the urban system and would thus pick the most optimal pairings regardless of their geographic proximity and physical distance between them. To embed the geographic information, we consider a complete graph where city units constitute the nodes and edges are weighted based on the Euclidean distance between the two city units.[2] This enables a selective trimming of the city pairs to be clustered based on a distance threshold such that only units or sets of units that are closer than the threshold are considered for clustering in Equation 6.1. Additionally, due to the agglomerative nature of

[2] Here, we use centroid-to-centroid distance where the centroids are obtained unweighted for city units polygons using the QGIS package. See Figure 6.3 for heat-maps showing the spatial distribution of these units.

Table 6.1: Summary of the boundary definitions used and the estimated exponents for A_n and Y.

Boundary	N_{min}	**No. of units**	**No. of units ($N > N_{min}$)**
C100	3895	2867	587
C350	7627	2928	481
C500	59,698	2475	104
C750	57,698	2021	112
C1000	55,031	1692	120
C1400	67,495	1435	97
C3500	66,671	859	49
LAU1	101,355	348	215
NUTS3	499,766	141	34
TTWA	510,149	173	28
URBAUD	159,581	83	55

While the constrained number of units has been used to estimate model parameters, for the administrative and functional economy boundaries the full set of units have been used in the hierarchical clustering.

such clustering approaches, an unconditional clustering would terminate only after having consumed all city units within a single unit. In order to both provide a termination criteria and an alternative benchmark for the clustering outcomes, we consider two parallel clustering procedures. In one, at each step we seek the city-pair with the smallest distance, $\eta_{(A+B)}$, in the other, in each step, we select the pair that also satisfies the added condition that its performance improves on both parent units. The clustering for both scenarios then terminates when the latter exhausts mutually improved pairings. In this way, we both limit the number of steps allowed to be taken in the original purely agglomerative approach and provide a clustering benchmark in which connections have improved on both units involved

For the implementation of the distance threshold, we consider two approaches. In the first, hereinafter denoted as CD, we choose a discrete distance threshold, trimming the graph of edges weighted over the chosen threshold and then applying the hierarchical clustering. In the second, denoted hereafter as SD, a more continuous setup is employed where a lower- and upper-bound for distance threshold and a step-size are selected. The graph is initially trimmed for the smaller threshold and the clustering algorithm is employed until all viable aggregations are exhausted. This is implemented as a node contraction where of the two original units to be merged the one with the smaller overall GVA is absorbed into the one with larger economic output, which consequently inherits the sum of the attributes of the two units. The threshold is then increased according to the step-size with some previously eliminated edges put back.

Pairing flowchart

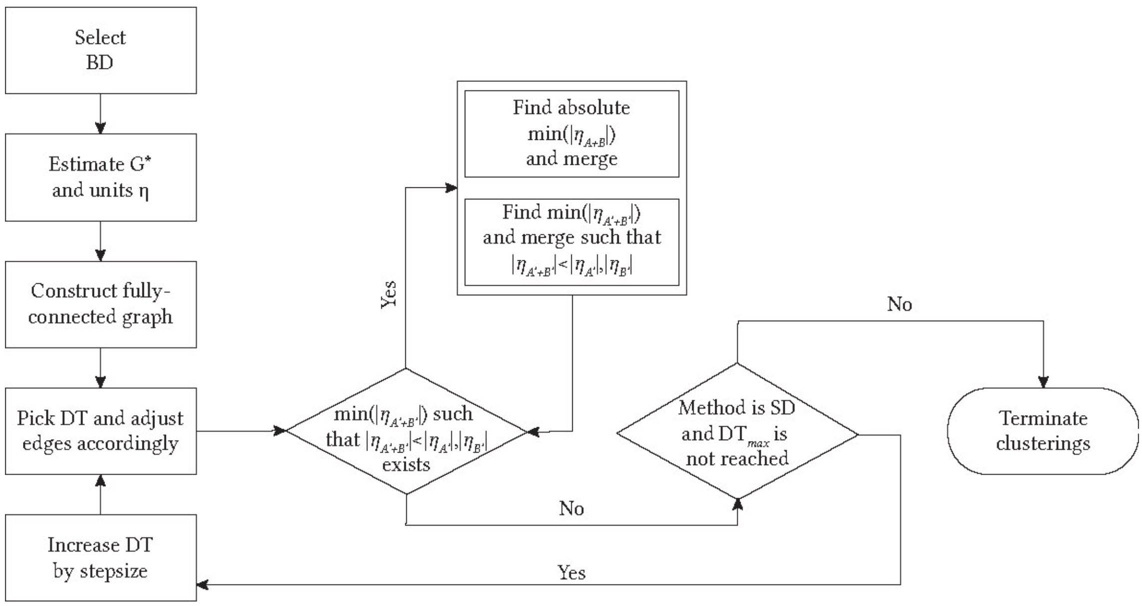


Figure 6.2: Flow chart capturing the process of hierarchical grouping of city units into city regions.

This is repeated until the distance exceeds the upper-bound specified. Together, the CD and SD methods enable examination of both city regions developed with no scale hierarchy and those developed prioritizing mobility starting from smaller local scales and moving up to larger regional scales. Figure 6.2 show simplified flow chart describing the overall process.

To isolate regional potentials, we also consider three regional scenarios. The base scenario, S0, is assigned as that with only the distance threshold limiting the clustering of city units. A second scenario, S1, is devised where, in addition to the distance threshold, city pairs with connections crossing the North-South divide are disallowed.[3] Similarly, a third scenario is considered, S2, regionally isolating the English North, South, and the Midlands according to the groupings of the NUTS1 areas, Figure 6.3. We implement the S1 scenario as a means to investigate pairings where the available units can be

[3] For S1, city units within each boundary are assigned a region based on their position relative to the North-South boundary developed by Dorling (2010).

Table 6.2: Summary of the scenario matrix and distance thresholds used.

		Clustering Approach		
		Purely Agglomerative		**Mutually Improving**
		S0	S1	S2
Linkage Method	SD	Starting at DT=20km expanding towards DT=180km with 10km step size		
	CD	Clustering at DT=20, 40, 60, 80, 100, 120, 140, 160, and 180km		

Pairing geographic scenarios

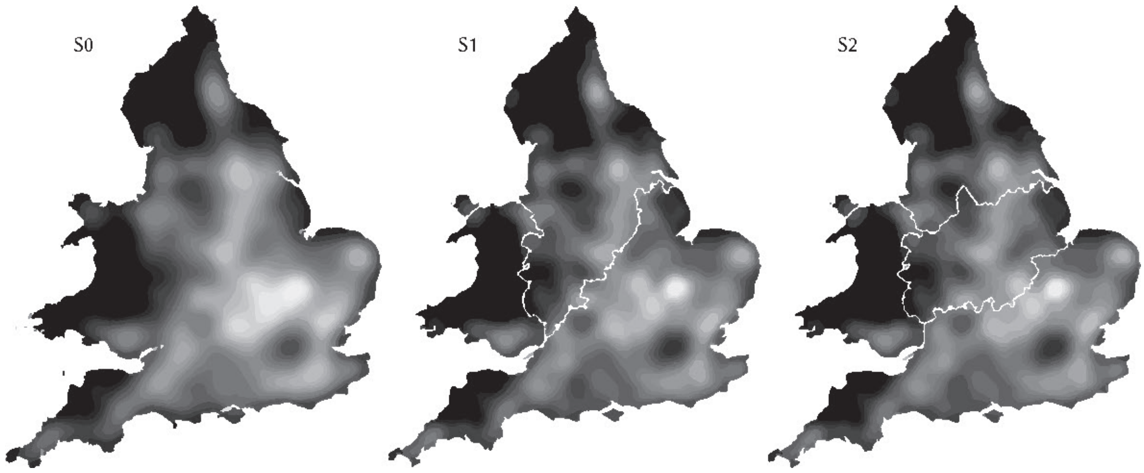


Figure 6.3: Maps showing the boundaries used for S0 (A), S1 (B), and S2 (C) with heat-maps showing overall distribution of all city units across all definitions.

considered to be more similar across a range of indicators, e.g. life expectancy to house prices (Dorling, 2010). This is while scenario S2 enables us to examine consistent alternatives to/for the current pattern of city-regions proposed in the north and the midlands based on LAU1 and NUTS3 units (Midlands Connect, 2017; Transport for the North, 2015). Table 6.2 provides summary of the scenario combinations, i.e. distance thresholds and step-sizes, considered in this study while Figure 6.3 shows the geographic boundaries used for scenarios S0-2.

6.3. City regions in England and Wales

We start by examining the resulting clusters for the local authority units (LAU1). Given that LAU1 units breakdown larger functional urban units, in particular that of Greater London Authority where a highly functioning inter-city transport system already exists, we would expect the clustering procedures, especially the SD scenarios, to capture these short distance internal pairings. This is tested for by mapping the LAU1 units to the TTWA units within which their centroids fall and then performing a frequency analysis on the occurrence of city pairings between TTWA units. Table 6.3 shows the top 5% of the most frequent pairings aggregated over all SD scenarios, i.e. combined S0, S1, and S2, for LAU1 units. As can be seen for the purely agglomerative approach, when mapped to TTWA units, the most frequent pairings ($f \geq 12$) do indeed show connections between units within the same TTWA, i.e. London, Manchester, and Derby, with the two most frequent capturing the connections within

Table 6.3: Showing the top 5% of the LAU1 pairings mapped to their parent TTWA with pair frequency.

Purely Agglomerative			Mutually Improving		
Origin	**Destination**	**Freq.**	**Origin**	**Destination**	**Freq.**
London	London	63	London	London	63
Slough & Heathrow	London	34	Slough & Heathrow	London	24
Manchester	Manchester	15	Leicester	Leicester	11
Slough & Heathrow	Slough & Heathrow	12	Medway	London	9
Derby	Derby	12	Brighton	Crawley	9
Chelmsford	Chelmsford	9	Manchester	Manchester	9
Chelmsford	Southend	9	Luton	London	9
Nottingham	Derby	9	London	Crawley	9
Birmingham	Worcester & Kidderminster	9	Leicester	Derby	8
Leicester	Leicester	9	Nottingham	Derby	7
Luton	London	9	Chelmsford	Colchester	6

London and between London and Heathrow as expected.[4] Moreover, 10% of all 214 mapped LAU1 city pairs are those capturing intra-TTWA connectivity and mobility. All the while, for the mutually improving approach, despite changes in the ranking of individual pairings, the overall mix of pairings shows very similar constituting members including mostly intra-TTWA pairings. While London already has an effective inter-city transport infrastructure managed through Transport for London (TfL) and Manchester is moving in that direction (Transport for Greater Manchester, 2017), the rest of these units are yet to implement such infrastructure systems flagging up a lack of adequate mobility provisions at spatial scales smaller than that of meso-scale regions. The important implication here is that intra-city projects targeting congestion, as they seem to be articulated currently, may be missing the broader problem of quality and diversity of available transport modes and the overall internal connectivity of urban areas.

[4] As we have previously noted, the original TTWA methodology does indeed aggregate London and Heathrow areas as the same TTWA for 2011 Census data. The final separation of the two areas is done based on results of stakeholder engagement and expert views (Coombes and Office for National Statistics, 2015).

Having sense checked the clustering approach, we proceed to examine the implications of city pair distance and choice of boundary on the city regions clustered.

6.3.1. Local versus regional

Figure 6.4 shows the cumulative distribution (CDF) of the distance between city units paired in each boundary definition disaggregated, in grey, for different geographic scenarios and distance threshold methods. The two red lines show the overall CDF of city-pair dis-

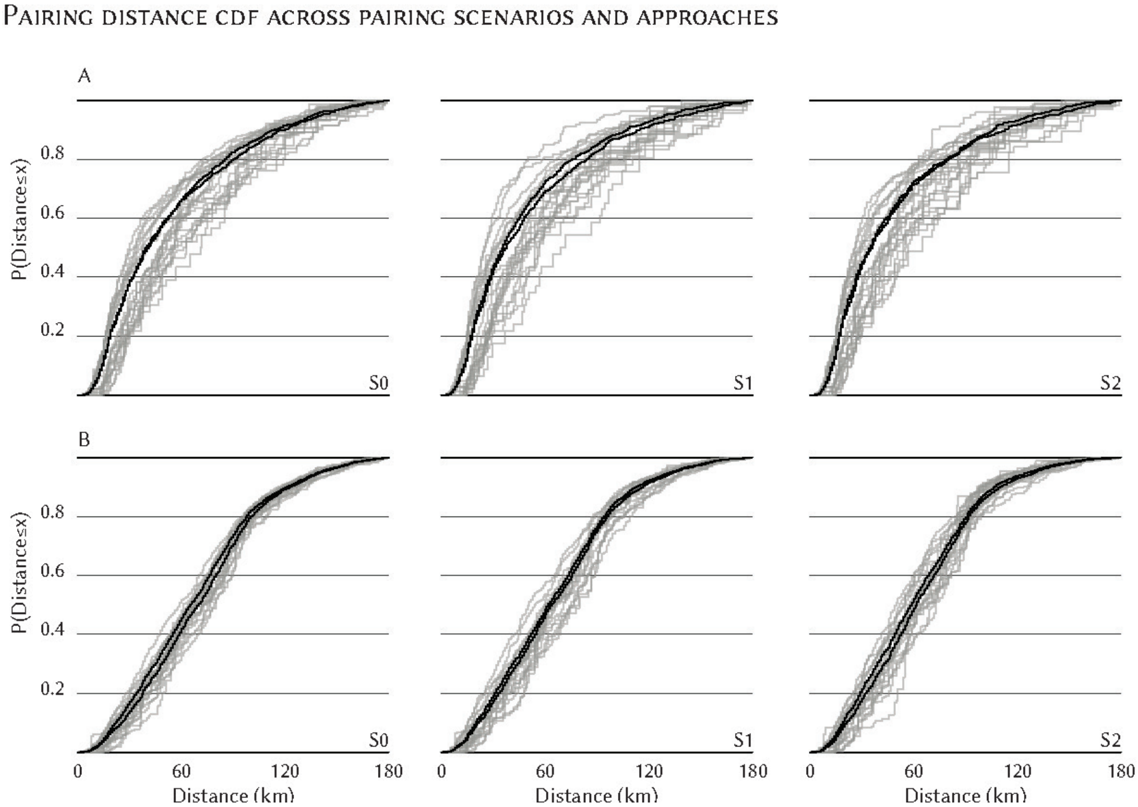


Figure 6.4: Cumulative distribution function of the distance between city-pair centroids in each geographic scenario for linkage methods SD (top row) and CD (bottom row).

tance across all scenarios and clustering approaches. It is quite clear that the choice of clustering approach, be it purely agglomerative or mutually improving, does not have noticeable effects on the distances over which potentially complementary city-pairs exist. Figures 6.5 and 6.6 additionally provide cumulative distribution of the city-pair distance disaggregated for three indicative distance thresholds at 60, 120, and 180 kilometers for CD and SD aggregation methods respectively.[5] Note that while the figures include distributions from both purely agglomerative and mutually improving approaches, there is no significant difference in the overall CDF of pair distance.

The noticeable difference between the distributions from the two methods is the more concave nature of the SD distributions compared with the more convex tendency of those of the CD method, Figure 6.5. This is mostly a result of the SD method exhausting local optimal pairings before moving up the distance threshold. Nevertheless, for the CD method, where there are no local distance prioritizations, the median pair distance grows an overall 77km from about 40km to 77km and then 117km for the most permissive scenario, S0, at 60km, 120km, and 180km thresholds respectively. The overall growth is only 52km for S2 from the same 40km median at 60km threshold. Distribution median for the SD method, however, grows from 26km to 35km and 41km for the same S0 scenario. Additionally, as can be

[5] We review distributions only at these three distance thresholds as these capture an appropriate range from smaller scales to mid-sized scales and larger ones with the remaining thresholds, already featured aggregated in Figure 6.4, fill in the pattern set between these three thresholds.

seen, this relative preference for shorter distance pairings remains more or less independent from the choice of boundary definition.

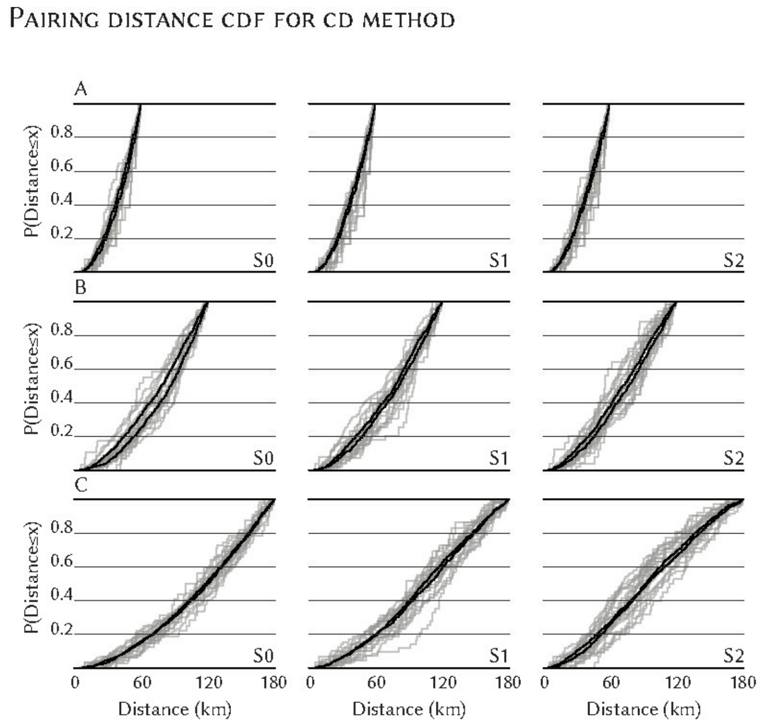


Figure 6.5: Cumulative distribution of the Euclidean distance between city pair centroids in each geographic scenario for linkage methods CD at distance thresholds 60km (A), 120km (B), and 180km (C). Note that overall distribution for purely agglomerative and mutually improving approaches are indistinguishable.

Considering the top 10% of the most frequent pairings in S0, S1, and S2 scenarios using the SD method, 21%, 27%, and 26% of all pairings across various boundary definitions take place between cities within the same TTWA unit. This prominence of short distance intra-urban solutions is also evident when we repeat the frequency analysis for the superposition of the clustering outcomes over all boundary definitions. While the SD method could be assumed partial towards shorter distances, the relative prominence and occurrence of within-city connections can be shown to persist even when considering clustering outcomes from the CD method at 180km distance threshold. Table 6.4 shows the percentage of intra-TTWA pairings comprising all pairings, the top 20%, and 5% most frequent pairings when mapping all SD and CD outputs to TTWAs and also those specifically of S0 scenario with CD method at 180km. We would have expected the intra-city median frequency to be smaller or to coincide with the overall distribution median were the intra-city pairings a small and insignificant part of the distribution or random occurrences within it. Despite the diversity of the city boundary definitions, distance thresholds, and clustering approaches, the frequency of intra-city pairings, however, remains of significance as

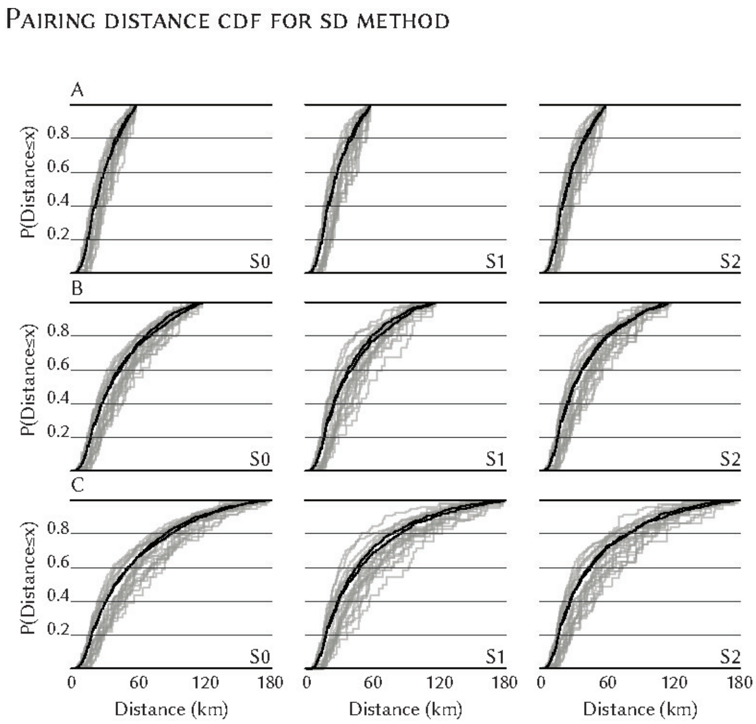


Figure 6.6: Cumulative distribution of the Euclidean distance between city pair centroids in each geographic scenario for linkage methods SD from the starting distance threshold up to thresholds at 60km (A), 120km (B), and 180km (C). Note that overall distribution for purely agglomerative and mutually improving approaches are indistinguishable.

Table 6.4: Percentage of intra-TTWA pairings across scenarios.

	% of intra-TTWA pairings					
	Purely Agglomerative			Mutually Improving		
% top pairing frequency	SD	CD	S0-CD180	SD	CD	S0-CD180
All	7.4	3.5	1.4	7.7	3.0	1.3
20%	18.9	7.3	1.6	16.2	5.9	1.0
5%	27.8	7.5	2.4	24.5	8.6	4.2

demonstrated by their larger medians compared with those of the overall frequency distribution even at largest distance threshold scenarios. It is also worth mentioning that the most frequent connection remains that of those connecting units within the London TTWA.

6.3.2.
City-regions and recurrent centers

More broadly, as a nationally driven infrastructure policy, the overall efficacy of agglomerative region building centered on the provision of mobility and transport infrastructure can also be explored by investigating the fraction of city units, out of total, the infrastructural and productivity shortcomings of which can be addressed through

Percentage of total cities paired in a city-region

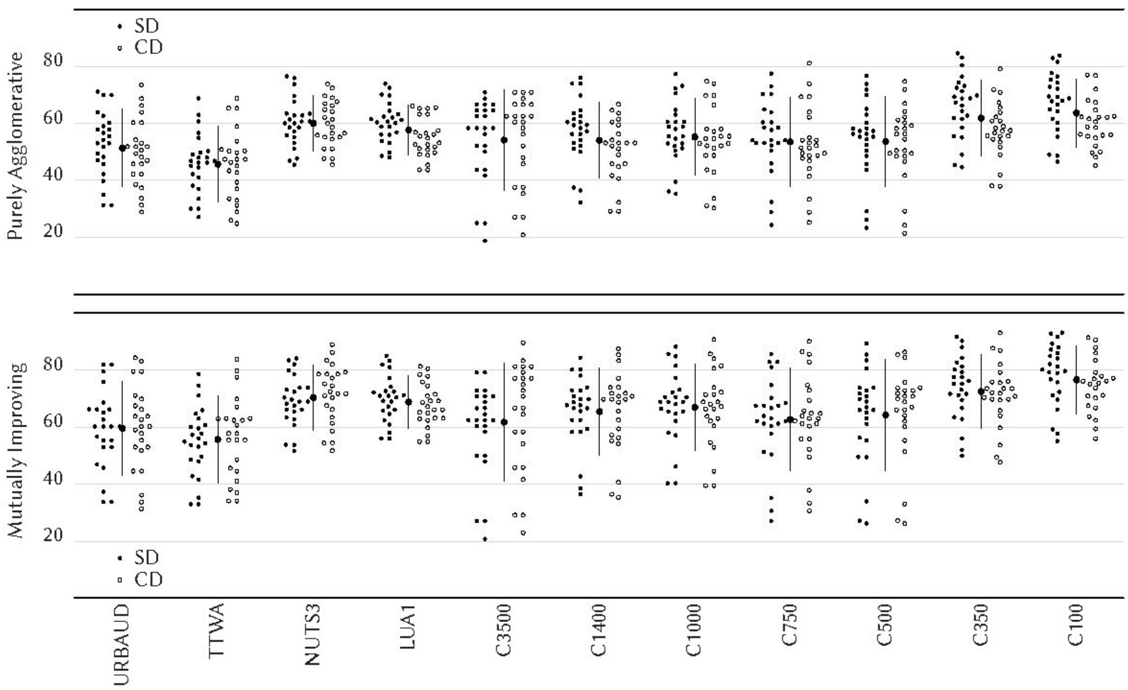


Figure 6.7: Strip-plots showing the distribution of the ratio of cities clustered in a city region over the total number of initial city units.

better connectivity with other city units. Figure 6.7 shows the strip-plot of this ratio calculated for each boundary definition using SD and CD methods for purely agglomerative and mutually improving approaches. Error bars show the standard deviation around the overall average ratio at each boundary definition regardless of the method used. As can be seen, the average ratios observed across boundaries hover more or less consistently around 60%.

The implications are twofold. First, considering administrative and functional boundaries, the inter-city transport connectivity as a way of addressing economic under-performance, at least in an English and Welsh context, does not appear to provide a universal solution. Despite few clustering outcomes reaching ratios as high as 80% towards the 180km DT, the average ratio remains around 60%. Spatial agglomeration arguments implemented through transport should, as such, be applied discerningly and wider national infrastructure planning needs to be tailored for a majority of city units individually across scales.

Second, the seemingly larger ratios of the density-based boundaries can be misleading and once again brings us back to the importance of intra-city connections laid out in the previous section. The administrative and functional economy boundaries, as compared with those that are density based, constitute a smaller number of overall units where each unit depending on the boundary might contain multiple urban cores and their hinterlands, the case of the functional economy boundaries, or vast extents of relatively low-density areas, the administrative boundaries. The density-based boundaries, on the other hand, could potentially break up such units into new ones around their populated centers, most of which while disconnected are close neighbors. These are then put back together through the clustering procedure when infrastructural needs are complementary.

Finally, we interrogate the geographic consistency and robustness of our synthetic city-regions. This is done by geographically embedding the aggregated TTWA-mapped frequency analysis as a weighted network where the weight of each edge is linked to the overall frequency of the connection between the two TTWAs or between units of other boundaries located within the two TTWAs.[6] Figure 6.8 shows this network visualization when aggregating across all scenarios (S0-2), methods (SD and CD), and distance thresholds isolating the top 1% of all edges.[7] The insets at the bottom show separate aggregations for SD only (A), CD only (B), and CD-180-S0 only (C). It should be noted that the 1% connected cluster in the North does not include Manchester and the edge is that of Bradford-Crewe. The partitioning shown has been done applying a modularity-based community detection algorithm finding communities where edge-weighted connectivity between community members is more significant than inter-community connectivity to the full extent of each graph (Blondel, Guillaume, Lambiotte, and Lefebvre, 2008). While the two main panels in Figure 6.8 show the most frequently recurring city regions regardless of the connectivity distance thresholds and/or regional reach and limit, the insets provide variations reflecting different planning priorities. Inset A, showing the most frequent links for the SD method, demonstrates city region configurations where intra-city mobility improvements have been prioritized. Inset B, in contrast, shows a multi-scale provision of connectivity effectively superimposing optimal pairings across scales, hence the larger connectivity. Lastly, inset C demonstrates a focus on long-distance pairings. It is noteworthy that community modularity for CD-180km-S0 broadly partitions units along Dorling's North-South divide (2010) used in scenario S1 while isolating London-Birmingham-Manchester as an individual community cluster. The London-Birmingham-Manchester grouping, especially the higher frequency London-Birmingham link, incidentally picks up the current major transport infrastructure project in the national pipeline (Infrastructure and Projects Authority, 2015).

[6] In this manner, each edge denotes an inter-TTWA pairing while self-loops denote intra-TTWA pairings.

[7] Note that the self-loop representing the intra-city connections within London has been removed in the highlighted edges for better scaling of the weighted sizes.

Of particular interest are, however, the differences and similarities of regional clusters created through the purely agglomerative and mutually improving approaches. Although the clusters produced by the two approaches are visually distinct, especially for those at CD-180km-S0, the combined optimal city-region of the Midlands centered around the Leicester-Nottingham-Coventry triad remains stable throughout. The only other high-frequency pairings to remain stable across approaches and scenarios are the intra-TTWA links within London and Manchester.

For intra-city transport at a TTWA scale, areas such as London, Medway, Cambridge, Chelmsford, Coventry, and Manchester show potential to benefit from an infrastructure that enables mixing within their TTWA boundary. Some of the same areas also constitute the larger urban areas at the core of larger city regions to be connected via inter-city transport schemes. For the most parts, when considering the overall network and insets A and B, the broader connected communities are consistent with a regional aggregation of NUTS1 areas. This is for the exception of the connectivity divide in the south of England between the south-west and south-east which is more consistent with the geography of the clusters developed by Arcaute et al. (2016) when analyzing the connectivity of the road network in Great Britain through hierarchical percolation. We should however note that a point to bear in mind regarding the intra-city self-loops is that while all these urban areas show a potential to benefit from a better-mobilized population within the boundary of their respective TTWAs, London is the only area currently equipped with an overall transport infrastructure capable of delivering this.

6.4. Chapter discussion

We begin the discussion with the acknowledgment of a common obstacle faced by spatial analyses of urban areas. Empirically, all spatial statistics, and scaling frameworks in general, are subject to the 'modifiable areal unit problem' (Openshaw, 1983). This is precisely why the approach presented in this volume has explicitly looked at realizations of city units at varying spatial scales and boundaries underpinned by a multi-scale hierarchical approach. By looking through a multi-level lens, we have, as stated, empirically examined the stability and consistency of the problem across spatial scales.

From an analytical perspective, by then mapping the clustering connections made to the TTWA units and examining connection frequencies, we have obtained persistent complementarities that remain stable despite changing spatial scales. Additionally, due to the intrinsic definition of TTWAs that implies areas within the same boundary

Pair-frequency network map

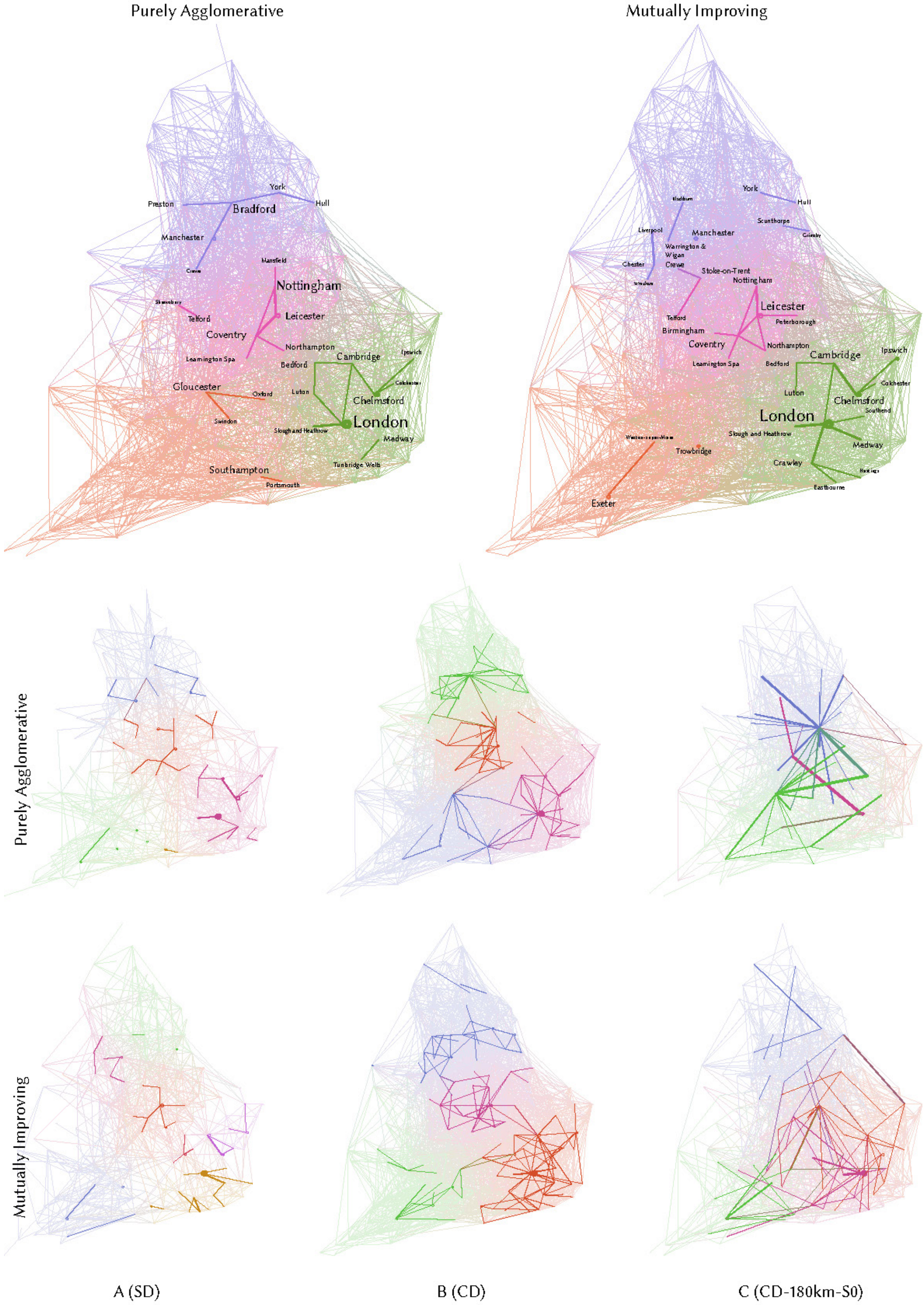


Figure 6.8: Weighted network of overall pair frequency highlighting the top 1%. Insets include pair frequency for SD method (A, top 5%), for CD method (B, top 5%), and CD-180km-S0 (C, top 5%) – label size is proportional to the city weighted degree.

constitute a unified economic marketplace, we can view intra-unit connections as existing complementarity within an existing urban unit that can be boosted through better intra-TTWA mobility, if not already in place similar to that of London. By contrast, the inter-unit connections then highlight currently competing units whereby complementarity exists such that were they to act cooperatively as a single and unified unit, given a mobility infrastructure enabling efficient inter-TTWA mobility, the larger metro area would achieve closer to optimal $Y - W$ performance.

Moreover, it is crucial to be aware that neither Bettencourt's model in itself or the clustering scenarios discussed here provide any recommendations on transport *investment* since the pairings are based on performance balance potential and not effective return on investment. The imbalance discussed is then not of transport per se but of a mobility-output trade-off. As an illustration, suppose one thinks of or expects each of city units at a given spatial scale to have an adequate economic performance balance on its own. At each boundary definition, then, there are two issues to consider:

1. is the overall output or urbanized area scaling exponent close to the theoretical, and
2. for each city is the estimate of G close to the theoretical optimal.

Note that the two are to some extent independent. An overall number of city units can show systemic mobility problems whereby the elasticities approach linearity while the $Y-W$ is optimal because they compensate for deviations in the scaling of one, say GVA, through deviations in the other. The clustering only addresses the potential for balancing $Y - W$ through matching complementary Gs.

We continue with a brief commentary on the long-term planning implications of using such scaling models for region building aimed at maximizing size-cost performance by an examination of the connections identified. A simple reading of social reactor model (SRM) used in interpreting these connections would frame the infrastructural intervention required as provision of better mobility. While generally a valid reading, interpreting all pairings without a consideration of the nature of the boundaries as transport related could prove short-sighted. When combining for a closer-to-optimal size-cost performance the model assumes an adequately mixing and mobile population. For contiguously urbanized areas, e.g. London and Manchester, intra-city connections indeed imply a need for an implementation of better transport infrastructure. For travel-to-work areas with a less uniform population extent and non-contiguous urbanized areas, e.g. Chelmsford, Cambridge, and Exeter, lack of adequate mobility is both a matter of access and the inherent distance between populated land patches. Considering Equation 3.3 again, $Y = \bar{g}a_0 l(N^2/A_n)$, a supposed recommendation for better intra-city mobility for such units would have to include both transport improvements, i.e. increasing average $a_0 l$, while also increasing effective population density through densification, i.e. decreasing overall A_n. This signals at a need for long-term densification of the most populated centers in these units.[8]

[8] This long-term significance of densification we have already discussed in chapter 5 when considering the continental comparisons drawn from the Randstad and Rhine-Ruhr metropolitan region.

A similar point can be raised about inter-city links where both units have similar conditions, e.g. Exeter-Yeovil, or those where one unit is significantly more uniformly dense and contiguous in urbanized area than the other, e.g. Bradford-Crewe and Coventry-Leamington Spa. In such cases, the clustering, as currently formulated here, recommends a pairing based on the *potential* that exists in the combined population size and urbanized area extent towards achieving agglomeration economies. The existing economic under-performance, however as seen in the previous chapter, results in the current clusters to have compensated for this productivity gap through the addition and increase of urbanized areas and hence population to maintain optimal size-cost balance. A more relevant interpretation of an increase in mobility and access for these scenarios would be policies aimed at further urbanization of the existing developed areas and moving inhabitants from several distant settlements into core contiguous urbanized areas over time. The *potential* population aspect of these pairings is then in line with the notion of urban 'borrowed size' (Alonso, 1973; Burger, Meijers, Hoogerbrugge, and Tresserra, 2015). From a purely cost-size perspective, however, such conurbations would benefit over time from densification and a decrease in the overall number of city units. Finally, from a scaling perspective, any policy, whether it be transport-related or not, proving successful in narrowing the economic under-performance needs to be accompanied by longer-term densification efforts in order not to result in escalating mobility costs over longer periods. This is true for low-density pairings in our clusters as well as units like London which can benefit from densification as a comparatively near ideal mobility infrastructure has already been implemented.

7

A General Discussion

Material prepared for this chapter have been used in the following–

Arbabi, H., Mayfield, M., Dabinett, G. (2019). Urban Performance at Different Boundaries in England and Wales through the Settlement Scaling Theory. *Reg Stud*, 53(6), 887-899.

Arbabi, H., Mayfield, M., McCann, P. (In Press). Productivity, Infrastructure, and Urban Density—an Allometric Comparison of Three European City-Regions across Scales. *J R Stat Soc A Stat.*

Arbabi, H., Mayfield, M., McCann, P. (In Press). On Development Logic of City-Regions: Inter- Versus Intra-City Mobility in England and Wales. *Spat Econ Anal.*

Arbabi, H., Mayfield, M. (2016). Urban and Rural—Population and Energy Consumption Dynamics in Local Authorities within England and Wales. *Buildings*, 6(3), 34.

The three rather stand-alone pieces of work presented thus far — throughout chapters 4 to 6 — form an inherently connected narrative exploring the effects of spatial scales on the observations of power-law scaling elasticities in relation with urban connectivity and economic output, Figure 7.1. In the present chapter, we see a succinct aggregation and synthesis of the collective observations and arguments from across these works as they relate to the broad questions outlined in chapter 1:

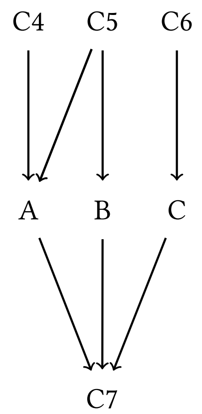


Figure 7.1: Schematic of the connection between chapters and research questions.

A. do the scaling arguments regarding connectivity and agglomeration remain valid across scales?
B. to what extent are the scaling insights transferable between different cities and systems of cities? and finally,
C. what are the effects of spatial scales and distance on the geographic patterns of connectivity-focused agglomeration within such scaling frameworks?

The rest of this chapter first provides a general discussion of issues concerning the results from previous chapters concluding with a dedicated discussion of particular policy insights for the English and Welsh case.

7.1. Spatial scales and population scaling of urban characteristics

In this section, we compartmentalize the discussion surrounding our results by initially summarizing the aggregated state of economic and areal scaling across the case studies from a Bettencourtean perspective and then addressing the quality and model-dependency of the observations.

7.1.1. Deviations of scaling exponents from theoretical expectations

As already stated, the overall case for better connectivity is often articulated from a stylized agglomeration perspective. From an urban scaling perspective, the economic output of an area is assumed as commensurate with the interactions between the urban population (Bettencourt, 2013; Glaeser, 2010). Consequently, an increased population size with increased connectivity will enhance economic output with the overall urban performance framed as the balance between increased interactions through improvements in connectivity and its associated costs addressable through urban densification.

Within a similar agglomeration paradigm, the multi-scale analysis of the urban system in England and Wales, as presented in chapter 4, does indeed signal at a systemic lack of theoretically assumed connectivity for a large portion of city units especially those located in the North or along the coasts. More broadly, however, this observation is not seen as particular to a specific spatial scale.[1] In fact, for the English and Welsh case study, this spatially systemic deficiency in connectivity, as evident from the closer-to-unity estimates of both β_Y and β_{A_n}, occurs at various spatial scales from very large city boundaries encompassing core urban areas and their wider hinterlands to those isolating uniformly dense urban cores. The analysis in chapter 5, examining the transferability of the framework's insight between urban systems, also details how these less-than-ideal elasticities appear to be unique to England and Wales and not an artifact of the spatial scales considered.

[1] We have, however, seen that better, i.e. more super-linear, regimes are observed for boundaries that enjoy more coherent provision of connectivity as a result of existing travel patterns, e.g. travel-to-work areas or OECD functional urban areas.

The observation of scaling exponents that are close to unity and cannot statistically be ruled out in favor of the theoretically expected sub- or super-linear regimes is not on its own unique in the context of urban indicators in England and Wales. Arcaute et al. (2015) do indeed report similar close-to-unity exponents for a variety of indicators[2] in England and Wales using output areas (OAs) as building blocks. While they then use this observation to argue that such deviations from the expected sub- and super-linear scaling is an indication that cities cannot be codified as simple power-law relationships, missing from their core argument is an assessment of whether such power-law models and formulations of cities, with Bettencourt's one amongst many, can explain such empirical deviations from their ideal theoretical expectation. What we have seen throughout chapters 4 and 5 is that not only the close-to-proportional scaling of both economic output and urbanized area predictable within social reactor model (SRM) framework but also meaningful given the context provided by the continental comparison in chapter 5.

[2] Unlike the work presented in this volume, Arcaute et al. (2015) focus on patent and employment counts and their adequacy as indicators for urban innovation (Bettencourt and West, 2010).

This brings us to the issues of the agreement — or lack thereof — of observed scalings with their theoretical expectation at particular spatial scales.

7.1.2. Consistency of scaling regime across spatial scales

Considering the population scaling of economic output and urbanized area, two issues needs addressing. These are an apparent dependence of exponent β_Y on density cut-off threshold [3] and interpretation difficulties involving the two administrative and statistical boundaries. Interpreting the seemingly increasing trend of β_Y with

[3] See Figure 5.2 in chapter 5.

the rising density cut-off threshold within the SRM is in fact both quite intuitive and trivial. Given the process followed in forming the city units using the density thresholds and the GEOSTAT grid cells,[4] increasing density cut-off values involve gradual trimming of peripheral grid cells of lower population density — note that the mostly peripheral nature of these cells is due to the expected auto-correlation and clustering of high density cells and the population density gradient away form these centers, Figure 7.2. As such, when comparing the exponents from, say, C1400 and C100, one is effectively comparing average connectivity within highly dense urban cores and across these cores and their hinterlands respectively. In this manner, the larger estimates of β_Y for increasingly larger density thresholds imply that the existing patterns of closer-to-theoretical mixing and connectivity are confined to the core high-density nuclei of city units which is congruous with one's expectation of intra-city availability of means facilitating connectivity.

[4] See *Data preparation and assembly* in chapter 3.

Figure 7.2: Schematic heat-map of population density showing its spatial distribution for England and Wales.

The final issue to address here is the noticeable divergence of the estimated exponents, both β_Y and β_{A_n}, for administrative and/or statistical boundaries, that is local administrative units level 1 (LAU1) and nomenclature des unités territoriales statistiques 3 (NUTS3), from the estimates at other spatial scales.[5] As alluded to briefly in chapter 3, administrative boundaries are potentially the least compatible with the internal model definition of a city unit. This largely happens as these boundaries can arbitrarily breakdown some urban areas down to smaller units while leaving others intact and hence provide for city units that are not necessarily coherent in terms of spatial scales. This is the case when comparing units of these boundary definitions within each country and also with those of the same definition from different urban systems. For example, Berlin with a population of nearly four million constitutes a single NUTS3 units while the Greater Manchester area of nearly three million is built up of 5 NUTS3 units. It is, however, of crucial importance to note that although these nation-wide and cross-national scale inconsistencies do cause the unexpected/unusual exponent estimates, the overall model formulation and comparison with the estimated boundary optimal remains consistent. This is to say that while from a utilitarian perspective these units do not provide for meaningful urban boundaries, if we insist on comparing urban economic and size-cost performance across these boundaries and attempt to push these units towards an individually obtained optimal performance then the social reactor model provides an assessment of the intervention required. The rationality of such an expectation remains a wholly separate issue and perhaps more dependent on the choice of individual unit at such boundaries.

[5] Appendix B provides a partial exploration of slight variations in input data and city unit boundaries and their effects on the estimated exponents.

7.1.3. Alternative models

The final point pertaining to the stability of observed scaling relationships and exponents and their interpretations across scales is issues relating to the choice of model. In chapter 3 we saw how accepting Bettencourt's four fundamental assumption frames the average-aggregate response of urban economic output in terms of the average connectivity of the inhabitants when deriving the power-law relationships empirically observed. We might, however, rightly question whether such abstractions of connectivity which has formed and shaped the arguments presented throughout chapters 4 to 6 are merely an artifact of the model chosen, i.e. social reactor model (SRM). Although the theoretical and parametric expectation's of the scaling exponents, β_Y and β_{A_n}, in SRM depend solely on the geometric and dimensional specifications of cities' geometry and their inhabitants' abstracted connectivity, D and H respectively, alternative derivations of such scaling exponents by others, while identifying differing parameters dictating the exponents' values, assign similar intuitions to these parameters. This is to say that these alternative scaling models, such as those developed by Gao et al. (2018), Ribeiro et al. (2017), or Yakubo et al. (2014), simply use alternative physical indicators in parameterizing mobility, access, and/or opportunity rather than arriving at a power-law scaling using fundamentally different principles.

As hinted at in chapter 3, these allometric urban models all share the core principle that socio-economic activity correlates with the inhabitants' connectivity and hence the means available to such models for tuning the value of the scaling exponent rest in mechanisms regulating the existence of such a link between individuals. Figure 7.3, for example, shows the variations of a generic super-linear exponent, β, against different formulations of mobility, accessibility, or opportunity in three alternative models to that of Bettencourt's.

In panel A, Gao et al. (2018) use α as a power-decay exponent to adjust the probability of two individuals connecting/meeting over ever-increasing distances starting with no restrictions at $\alpha = 0$.[6] In panel B, Ribeiro et al. (2017) use a formulation where the fractal dimension of the city, D_f, is combined with a given distance-decay parameter for the influence of interactions over distance similar to α in panel A.[7] Similarly, in panel C, m fulfills the role of the distance-decay parameter articulated by Yakubo et al. (2014) as geographical constraints inhibiting inhabitants' mobility and hence number of encounters over increasing distances.

[6] Note that the $\beta = 2$ at no distance restriction essentially reproduces Bettencourt's crude count of total possible city-wide encounters as $N \times (N-1) \approx N^2$.

[7] The combination of D_f, as estimated through the box-counting method, and a distance-decay exponent works similar to Bettencourt's D and H in estimating overall number of human interactions and the sum total of their outcomes.

For the sake of completeness, we have to note that a number of these models (Gomez-Lievano et al., 2017; Yakubo et al., 2014) also utilize individuals' demographic characteristics when considering probability of a given interaction between a pair. However, the recurring factor across all models which determines and explains the less-super-linear-than-expected scaling is the one controlling the formation of encounters over distance and hence overall urban connectivity. As such, although we have used Bettencourt's particular formulation since it is the only model readily providing a size-cost function, a multi-scale examination of scaling observations in urban networks using any other model of this family is expected to point towards similar connectivity issues.

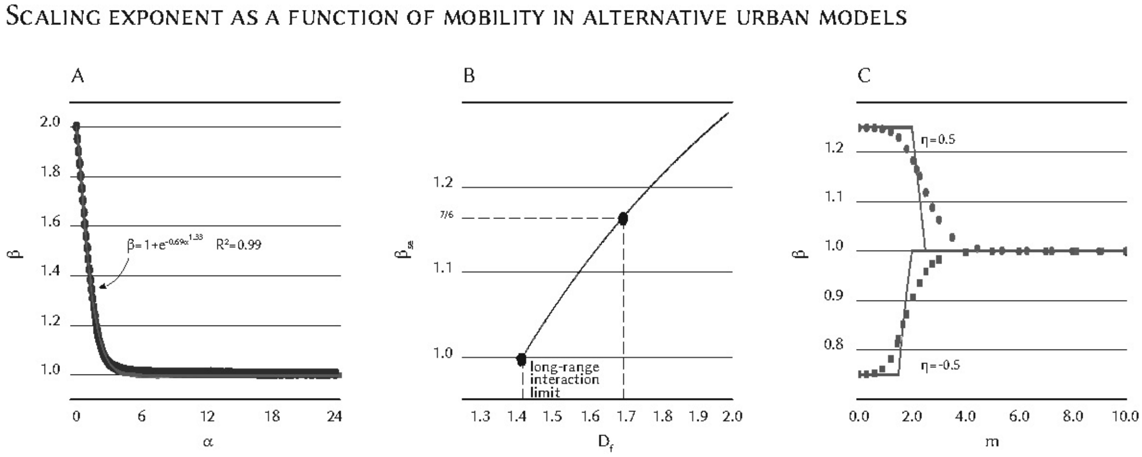


Figure 7.3: Variations of a generic super-linear exponent as a function of model parameter regulating individuals' connectivity for models developed by (A) Gao et al. (2018), (B) Ribeiro et al. (2017), and Yakubo et al. (2014).

7.2. Insights for the English and Welsh case

To investigate the effect of spatial scales on connectivity-driven agglomeration elasticities, the overall work presented in this volume has taken advantage of the fairly unique context provided by the urban system in England and Wales (EW) due to its regional economic imbalance as set out in chapter 4. As repeatedly pointed out throughout preceding chapters, the urban network in England and Wales has long exhibited and continues to exhibit an extreme inter-regional divide and imbalance with regard to economic output, productivity, and a variety of other socio-economic indicators (Dorling, 2010; McCann, 2016; Rowthorn, 2010).

The most recent assortment of mitigating measures devised to remedy this inter-regional divide has come to include provisions of inter-city, and to a lesser degree intra-city, mobility infrastructure (HM Treasury, 2017; Osborne, 2014; Transport for the North, 2015).

These have been argued for by referencing the body of empirical observations pertaining to the elasticities of economic output and productivity with regard to city size and density (Ciccone and Hall, 1996; Rosenthal and Strange, 2004). Following such empirical observations and relying on stylized agglomeration reasoning, the current wave of infrastructure programs are argued to tackle the output and productivity gap by enabling neighboring cities in the North to behave as a virtual urban unit of a larger size.

In this section, we review a collection of discussion points on the planning insights derived from Bettencourt's social reactor model applicable to the English and Welsh urban network system. We preface this by noting that the arguments offered in the following provide an isolated set of insights from an urban scaling perspective alone. With regard to connectivity-led attempts at achieving theoretically expected scaling economic output, these arguments would lack considerations of practicality and/or plausibility of implementation given concurrent factors such as dynamics governing the housing market (Cheshire, Hilber, and Kaplanis, 2015), infrastructure finance (P. O'Brien and Pike, 2015), and/or effectiveness of transport infrastructure as a driver of productivity (Crescenzi and Rodríguez-Pose, 2012) compared with increased connectivity as abstracted in an urban scaling framework.

7.2.1. Is it a connectivity problem?

Our observations in chapter 4 point to an alternative facet and/or framing of the North-South divide in England and Wales. The broader socio-economic division between the two regions can be reformulated from a size-cost performance perspective and presented in terms of long-term planning needs. We have seen that despite the higher-than-expected economic output, according to the SRM, economic success in a majority of the South-East appears to be achieved through mounting mobility costs as compared with idealized urban cities of the same population that would have exhibited smaller urbanized extents. This examination of Bettencourt's explicit formulation of the size-cost performance balance shows that the cities in the South-East, and London particularly, have in fact grown too large and require built-up area densification. Put in a broader context of the literature, this is, however, in contrast with Cheshire's (2013) recent criticism of densification and urban containment strategies labeling them theoretically grounded but without empirical grounds. Nevertheless, Arcaute et al. (2016) use percolation at different distances on the UK road network to obtain a hierarchal classification of the road transport network and by extension the clustering of the geographical regions as represented by their road connectivity where the England and Wales network initially collapses into one radiating out from London connecting the southern regions and the other connecting the North, Wales, and Cornwall. We observe a similar pattern to some extent where the South-East, regardless of the spatial scales considered, exhibits an overwhelming need for more compactness, especially along the radiating motorways, while the majority of the North suffers from poor intra-city connectivity, Figure 7.4.

North-south divide

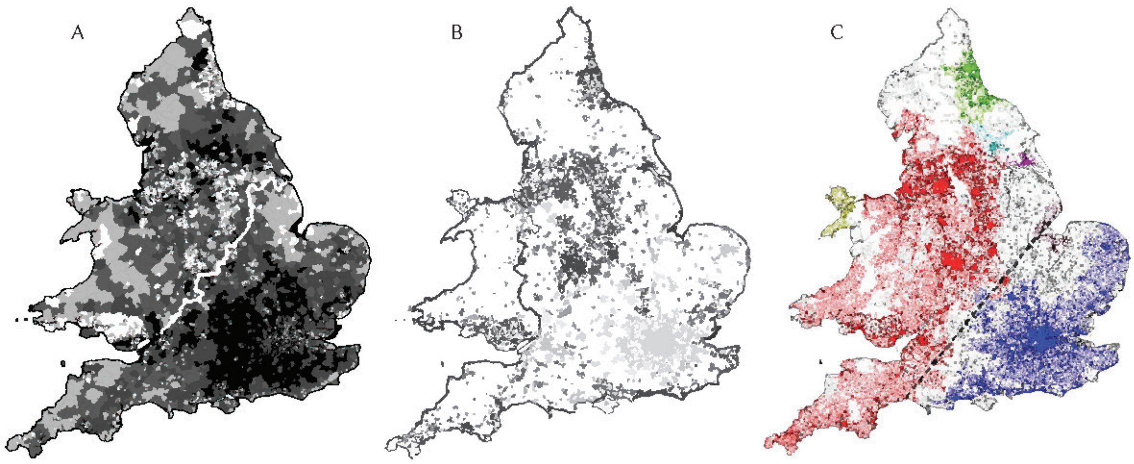


Figure 7.4: Maps of England & Wales showing (A) median household net weekly income in 2011 with Dorling's (2010) speculative dividing line overlaid, (B) estimation of η for urban units at $100^{prs}/_{km^2}$ density threshold (C100), and (C) clustering of road network at a percolation distance threshold of 740m with each color denoting a separate connected component (Arcaute et al., 2016).

As such, the examination of the English and Welsh urban system from a scaling perspective does, in principle, point towards inadequate level of human interactions and hence connectivity limitations as contributing towards the disparity in the regional economic performance balance across England and Wales. This is the point at which the discussion is brought back to the matter of the suitability of better mobility and transport infrastructure strategies borrowed from regions like the Randstad or Rhine-Ruhr metropolitan region. The literature arguing the suitability and efficiency of the poly-centric planning is not lacking. However, these most often study regions the core centers of which have developed an overall complementarity in terms of function, both economic and infrastructural, over time (Meijers, 2005) and as we have seen in chapter 5, the continental differences extend beyond those relating to particular transport infrastructure differences. It is of importance to note that this is in no way an argument against the implementation and investment in an efficient transport infrastructure increasing connectivity, mobility, and economic opportunities across the North of England but rather

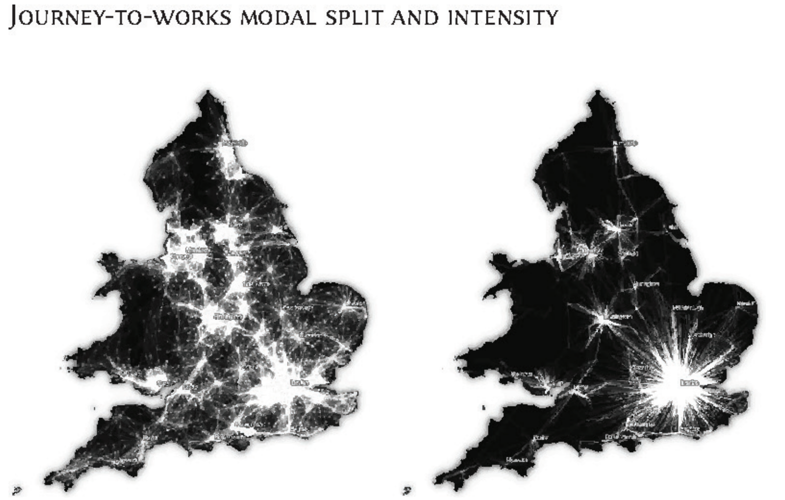


Figure 7.5: Network schematic of journey to work by car (A) and rail (B) in 2011 filtered for journey counts larger than 4 (Rae, 2016).

one pointing out that although such infrastructure would perhaps boost economic output over time, the issue of overall performance balance of cities would appear to require more than simple single-system policies especially when borrowed from other regions the effective similarity of which with the target regions of these policies are not explored thoroughly. These, however, bring up issues regarding the choice of transport type and mode.

Bly et al. (1980) note that the gap between overall generalized cost[8] of transport via private vehicles is so considerably smaller than those of the public transport, specifically buses, that even a full subsidization of fares would be unlikely to attract a significant number of motorists. They also suggest that passenger rail services addressing the travel-to-work commuting journeys in large urban conurbations and their up-take are vulnerable, in long-term, to the small window during which households tend to move closer to their primary place of employment potentially leaving the rail services obsolete long-term. This is of particular interest when considering the proposed inter-city rail upgrades in the North of England. This is particularly problematic when one considers the journey-to-works in England and Wales by transport mode.[9] Figure 7.5 shows the geography of journey-to-work in England and Wales for the year 2011 highlighting preference towards cars in the North as compared with rail in the South-East. In its absurd extremes, under such a paradigm, a faster connection between, say, Leeds and Manchester, or more appropriately Sheffield and Leeds, is of any importance only in the initial period after which those empowered by it to find employment in the other city relocate, undermining the function of the service long-term. Additionally, when comparing the historic trends of public transport operating cost and subsidies among 15 countries, the UK exhibits similar operating

[8] This, the authors define as ‘time and money of traveling by car’ (Bly, Webster, and Pounds, 1980).

[9] Although Bly et al. (1980) is a dated source, Nash et al. (2019) show a similar trend between 1995-2014, especially with regard to subsidies and prices, in their comparative analysis of Germany, France, Sweden, and the UK.

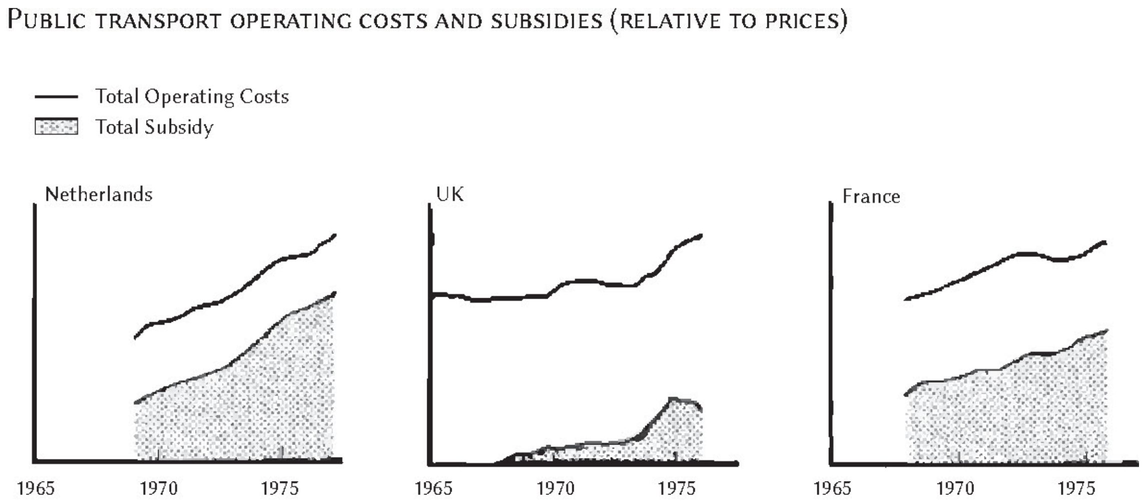


Figure 7.6: Historic trends in operating costs and subsidies, relative to 1976 prices, for the UK, the Netherlands, and France — reproduced from Bly et al. (1980).

costs to countries such as the the Netherlands and France while spending a fraction of what they do subsidizing services, Figure 7.6.

Additionally, when considering the share of transport modes across several countries, Pucher and Buehler (2008) observe that the UK, unlike its European counterparts, shows patterns more similar to the US and Australia most evident in cycling constituting only 1% of all journeys, suggesting that the US and UK have facilitated the dominance of private cars not just as a result of their transport and infrastructural policies but through a combination of land-use, planning, housing, and taxation policies. This is in stark contrast to trends in the the Netherlands (27%),[10] Denmark (18%), and Germany (10%) where the take-up of bicycles seem to remain stable regardless of individual characteristics such as gender, age, income, etc. The authors go further to attribute these differences to the larger portion of shorter-distance journeys in the European countries as opposed to those in the US and UK.

[10] Handy et al. put this at a whopping 37% for Amsterdam. Pucher and Buehler (2008) attribute this to an overhaul of urban planning policy during the 1970s. They also note that all three countries have invested consistently in bicycle infrastructure and facilities and maintain national master plans for bicycling.

A final point of importance to note is hence the subtle nuance required in interpreting issues of inadequate connectivity. Strictly speaking, from the perspective of the social reactor model, there is no theoretically prescribed preference for the method through which these mobility improvements are to be achieved. That is to say, the model's perceived lack of adequate mobility is formulated in terms of the geometry of inhabitants' average-aggregate area of influence and that smaller than desired values could be a results of various factors ranging from the absence of adequate physical infrastructure to expensive modes of transport. In such a paradigm, planning measures attempting at addressing a lack of adequate connectivity need to exercise care regarding the spatial scales over which connectivity issues are being mitigated.

7.2.2. Is it an inter- or intra-city problem?

The importance of an explicit consideration of a spectrum of spatial scales is twofold. The first concerns the common obstacle of the 'modifiable areal unit problem' (Openshaw, 1983). Explicitly considering realizations of city units at varying spatial scales and boundaries underpinned by a multi-scale hierarchical approach enables an empirical examination of the stability and consistency of the problem across spatial scales. It can then be argued more generally that limiting the spatial scale of infrastructural analysis and intervention only arbitrarily constraints available solutions for a problem that otherwise appears to require a more concurrent consideration across spatial scales.

The second, as mentioned in the last section involves the inherent differences in the manifestation of economic performance and mobility problems at different scales. Although economic under-performance is more easily noticeable at larger inter-city scales,[11] the mobility and performance balance problems in England and Wales appear to be persistent across scales and within intra-city boundaries. As explored in chapter 6, the distribution of distance between potential city pairings suggests a mobility problem that is both persistent and addressable at smaller scales and within intra-city boundaries. Consequently, because of the inherent hierarchical nature of spatial scales and distances, although inter-city transport-led agglomeration strategies are fitting, when implemented alone, would only mask transport and mobility shortcomings at smaller scales without addressing underlying causes of such under-performance. Diao et al. study of the inter-city high speed rail in China and its negative effects on intra-city congestion provides a demonstration for this point (2017). Meanwhile, initially addressing the $Y - W$ balance at smaller scales and distances would inherently be beneficial to larger scale connectivity. This would enable the transport infrastructure implemented over larger distances to contribute towards homogeneously increasing the urban system's overall baseline productivity. In contrast, a larger-distances-first priority would still be at the mercy of inadequate connections or overwhelming mobility costs at smaller scales. As such, although the currently planned rail heavy infrastructure policy would provide for the inter-city mobility needs, it would not address needs for better mobility, connectivity, and accessibility at other spatial scales.

[11] As previously mentioned, this is partly due to the limited availability of data to only certain functionally arbitrary administrative or statistical boundaries.

7.2.3. Long-term multi-scale implications

Last to address is the city regions clusterings explored in chapter 6 and the longer-term needs of the urban system in England and Wales in terms of overall performance balance. The clustering exercise in chapter 6 includes two important observations. Firstly, city unit agglomeration opportunities occur and are frequent within existing travel-to-work area (TTWA) boundaries or at relatively short distances. Secondly, a preference for optimal performance balance would demand long-term densification of currently rural areas. These findings are largely independent from city boundary definitions. It should, however, be highlighted again that the city pairings explored in this work are solely based on a performance balance potential. It is crucial to be aware that neither Bettencourt's model in itself or the clustering scenarios discussed chapter 6 provide any recommendations on transport *investment* since the pairings are not based on effective return on investment and as such combinations of city units in regions assembled prioritizing size-cost performance balance are not necessarily in agreement with those advocated by political agendas. The imbalances discussed is then not of transport per se but of a mobility-output trade-off. As such the observation of recurring short-distance and intra-unit pairings is indicative of more *opportunities*[12] at smaller scales than typically formulated through projects such as High Speed and bodies such as the Transport for the North.

[12] These opportunities can be thought of as starting scenario to be used as benchmarks or complement existing micro-scale land-use and transport modeling tools.

As alternative means to increase UK economic prosperity in a European context, Bettencourt and Lobo (2016) also explore the policy of moving the population away from small settlements to populate cities of medium size, i.e. those with profiles similar to key cities of the Northern Powerhouse, which remains in agreement with the findings presented in this work, especially those from chapter 5. In reality, the required strategy as far as the urban local administrative units level 1s of the Northern Powerhouse are concerned is potentially a mixture of both long-term densification in each region designed to scale back the extent to which the built pattern in local authorities such as Wakefield or Barnsley have sprawled and concurrently enhancing mobility, the transport infrastructure, and its efficiency to boost and strengthen the baseline economic output. Such arguments, while on the surface would appear mostly compatible with those promoting brown-field development in an English context, might not be consistent with their objectives in practice. This is due to potential discrepancies between places with brown-field space available for development and those where densification strategies are needed as indicated by Bettencourt's model (McCann, 2016).

As for the recurring densification-related observations, while the limited energy demand-related examination of scaling patterns in cities could reinforce a preference for higher densities when considering economies of scale, caution must be practiced in using this understanding for policy purposes.[13] The limited evidence mostly suggests that there are overall savings in terms of total energy consumption associated with higher density urban settings. This on face value could lead to simple advocacy for a preference in higher density developments. It, however, should be noted that at least for the network of cities in England and Wales, despite the clear existence of these trends, the practical savings may not be worth other potential technical and socioeconomic expenses (Echenique, Hargreaves, Mitchell, and Namdeo, 2012) as each 1% increase in population density at LAU1 boundaries appears to only result in approximately 0.3% and 0.06% decreases in per capita transport and domestic electricity consumption, respectively.

[13] See Appendix A for an examination of the broader trends of energy consumption versus urban population and density emphasizing potential energetic efficiencies associated with high-density urban living in England and Wales.

Finally, we would be remiss, however, if we did not also review the remaining shortcomings. The majority of models from the same family of the one used here start from the assumption that the units under study are in fact uniformly urban and functional economic catchments (Bettencourt, 2013; Gomez-Lievano et al., 2017; Yakubo et al., 2014). Unfortunately, this leaves them highly sensitive to the urban population count at each spatial scale and hence the choice of boundary used in that scale (Arcaute et al., 2015). Although it should be noted that while such fluctuations were observed in the G estimates from the model used in this study when considering slightly different boundaries at similar scales, the determination of the planning needs relative to their idealized counterparts remains more or less consistent.[14] Louf and Barthelemy (2014) argue that until a comprehensive, universal, and mechanistic understanding of how cities are formed, evolve, and function is developed such models should not be used in shaping policy advice. However, given that planning policy will be formed one way or the other and that the current economic agglomeration models informing policy not only suffer from the same fundamental lack of universality but are also as previously mentioned placeless and single-scale in nature, providing and considering alternative complementary pictures of city performance and infrastructural needs would benefit the overall policy and planning debate.

[14] Appendix B provides an examination of such potential variations.

7.2.4. *A speculative discussion of connected and autonomous vehicles*

Before concluding this chapter, it is interesting to review a remaining peculiarity of the current transport infrastructure plans for the North of England.[15] As discussed, although a rail heavy infrastructure policy would provide for the inter-city mobility needs, it intrinsically cannot address needs for better mobility, connectivity, and accessibility at finer spatial scales. More importantly, while a multi-scale approach to the provision of mobility, as we have seen, is fundamentally supported from the perspective of agglomeration theory and an assortment of urban scaling models, there is no theoretically prescribed preference for the method through which these mobility improvements are to be achieved. The rail-heavy focus of the current projects in England and Wales, especially those of the Northern Powerhouse and Midlands Connect, therefore appear to be due to the influence of continental comparisons drawn mainly from the rail infrastructure connecting the Dutch Randstad more than anything else (Swinney, 2016; Transport for the North, 2015). Meanwhile, current predictions put the widespread use of connected and autonomous vehicles in the UK at some point during the 2020s arguably prior to the first planned High Speed 2 service scheduled to take place between 2026-2033 (House of Lords, 2017). This brings about an ancillary question: would disruptive technologies such as connected and autonomous vehicles create more operationally and constructionally convenient alternatives to passenger rail infrastructure for providing and improving mobility and accessibility across spatial scales?

[15] The arguments here are beyond the immediate scope of this volume and are presented as a speculative critique of rail-heavy transport interventions in North of england.

Work journeys in the UK are already modally dominated by cars (Department for Transport, 2017). Figure 7.7 shows the modal split of work journeys in England and Wales, highlighting the prominence of road-based journeys in across the Midlands and the North, accompanied with the most frequent city-pairings developed in chapter 6. In a scenario already dominated by cars and hence the road infrastructure, the concurrently scheduled and budgeted road performance improvements in terms of congestion reduction and journey comfort are likely to move traveler choice even further towards cars and push upwards the acceptable distance to be covered through private mobility (HM Treasury, 2017). An analysis carried out by Yap et al. (2016) puts forward the claim that connected and autonomous vehicle would not only replace regular vehicles but are in fact expanding their role so to overtake the ecological niche in which conventional inter-city and public rail transport has developed. Additionally, they show that using connected and autonomous vehicles as a short distance mode of transport, e.g. from train station to final destination in town, is not an appealing perspectives for potential customers. It is indeed that such a use would deprive the CAVs from one of their most attractive features, which is the possibility of achieving single-mode journey over long distance without the discomfort of driving through them. In this way, CAVs represent the future of local

JOURNEY-TO-WORK MODAL SPLIT AND CITY-REGIONS

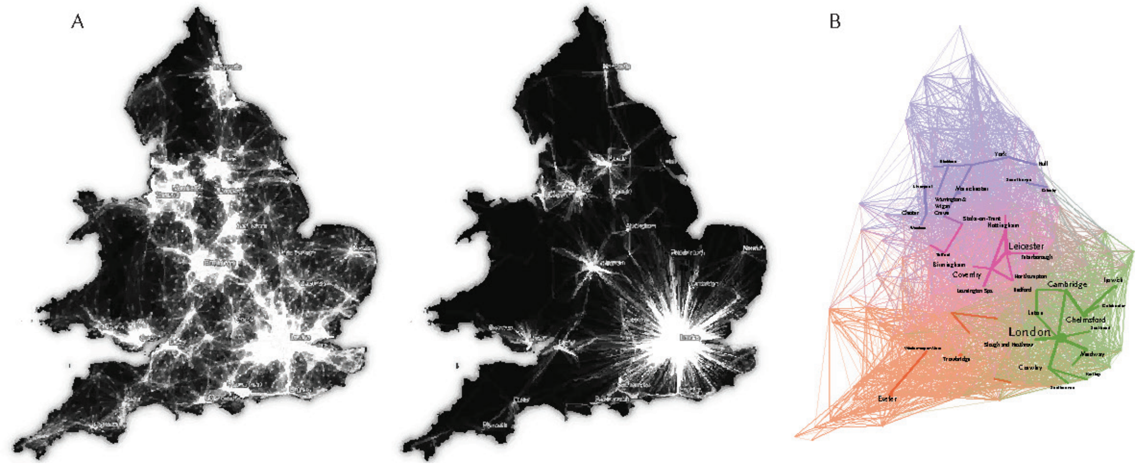


Figure 7.7: Maps of England & Wales showing (A) road-rail (left-right) modal split of work journeys in 2011 (Rae, 2016) and (B) the city unit clusters developed in chapter 6.

commuting as much as that of medium range travel.

Finally, there is the inherent difference in the topology of the infrastructure supporting connected and autonomous vehicles and passenger rail. In the case of inter-city transport, the rail network, by design, provides node-to-node movement between stations in different cities. As such, rail infrastructure can only comprise a single mode of multi-modal travel-to-work journeys requiring alternative modes, private or public, at either or both origin and destination stations. This means isolated improvements and upgrades of passenger rail network simply increase nodal pressure on other modes of mobility increasing capacity and frequency of rail services at the expense of cascading congestion over nodes they share with other transport modes. Potential road congestion improvements available through a socially optimized routing (Youn, Gastner, and Jeong, 2008; J. Zhang, Pourazarm, Cassandras, and Paschalidis, 2016), on the other hand, enable better mobility without a change in overall flow intensity over the road network through connected and autonomous vehicles. This in conjunction with the fact that road network provides the underlying infrastructure for multiple transport modes means that connected and autonomous vehicles are capable of side-stepping the nodal cascading pressure problem entirely as the connectedness applies across private and public vehicles operating on the infrastructure.

8

Conclusions

Material prepared for this chapter have been used in the following–

Arbabi, H., Mayfield, M., Dabinett, G. (2019). Urban Performance at Different Boundaries in England and Wales through the Settlement Scaling Theory. *Reg Stud*, 53(6), 887-899.

Arbabi, H., Mayfield, M., McCann, P. (In Press). Productivity, Infrastructure, and Urban Density—an Allometric Comparison of Three European City-Regions across Scales. *J R Stat Soc A Stat.*

Arbabi, H., Mayfield, M., McCann, P. (In Press). On Development Logic of City-Regions: Inter- Versus Intra-City Mobility in England and Wales. *Spat Econ Anal.*

Arbabi, H., Mayfield, M. (2016). Urban and Rural—Population and Energy Consumption Dynamics in Local Authorities within England and Wales. *Buildings*, 6(3), 34.

Throughout this work, we have focused on the subtleties that impact the generality of connectivity-based agglomeration arguments and the bigger-does-more-with-less principle from a multi-scale perspective. The main original contribution of this work thus lies in its examination of this often-overlooked aspect of connectivity-focused spatial agglomeration which is the choice of spatial scales and its effect on the observation of power-law scaling exponents. An additional contribution of this work rests in its use of urban scaling models which enabled an assessment of optimality in the trade-off between economic output and mobility costs accounting for connectivity and ease of access within cities coupled with their built density. Furthermore, to enable an examination of the effects of spatial scales on the geographic patterns of connectivity-led agglomeration, we combine Bettencourt's allometric urban model with a hierarchical clustering algorithm to offer a novel and mathematically grounded means of constructing city-regions based on urban size-cost performance balance at various spatial scales and distances.

The adoption of an urban scaling framework in this work has provided a number of advantages when compared with the approaches of the existing literature. Particularly, in terms of practical applications, the framework's significantly more parsimonious nature has enabled a study under circumstances where data available is relatively sparse. Moreover, the few fundamental assumptions underpinning the model used in this work are more general and driven by empirically observable average-aggregate behavioral patterns of cities and the urban systems to which they belong. Finally, due to the allometric framework's roots in the physics of self-organizing systems, such models provide a direct link to the rapidly growing area of complexity theory enabling such formulations of cities to maintain compatibility with others of such nature that focus on other aspects of cities besides economic-energetic performance.

8.1. Summary of empirical findings

In chapter 4, we initially reviewed new insights on the effects of spatial scales on urban performance balance. Using the urban system in England and Wales as an empirical testbed, we also explored the extent to which urban performance balance is influenced by the connectivity across scales, i.e. from intra-city to inter-city. The framework provides an explicit formulation and hence enabled an examination of the balance between economic output, connectivity, and associated mobility costs incurred with reference to the actual physical extent of cities at various spatial scales. In the specific case of the urban network of England and Wales, the analysis signaled at a systemic lack of adequate connectivity for a significant portion of city units considered at various spatial scales. More broadly, although effects of urban connectivity on economic performance are more easily noticeable at larger inter-city scales, the results suggest that connectivity issues tend to remain recurring at smaller scales and intra-city boundaries. These findings are also largely independent from city boundary definitions.

Following the isolated multi-scale examination of urban performance in a single urban system, we then investigated the universal transferability and applicability of such agglomeration-based connectivity-driven measures between and across different urban systems. In doing so, chapter 5 outlined a comparative analysis of scaling patterns in three European city-regions, i.e. the English North, Randstad, and Rhine-Ruhr metropolitan region, and the wider urban systems to which they belong across multiple geographic scales. The results demonstrated that although inter-city connectivity arguments can be considered transferable when comparing overall national performance of urban networks, a spatially multi-scale examination of the needs for better connectivity and/or densification in a given urban system can be made in isolation without requiring external comparisons. This, at a first glance, may appear to paint such regional comparisons trivial. However, the results also show how these regional comparative approaches are essential in identifying certain nuances which cannot be identified by looking at single-case data. Indeed, the multi-scale examination of the regional deviations from ideal scaling expectations of economic output and urbanized area reveal a general trend of observing larger-than-expected economic output with more-than-expected population densities regardless of the choice of spatial scale or city-region boundary pointing towards a deeper interplay of productivity, urban connectivity, and density. This essentially suggests that if connectivity-driven interventions do not drive and/or are not implemented in tandem with urban densification then they may not be likely to deliver the intended productivity gains on their own.

Finally, in chapter 6, we explored the effects of spatial scales and distance on the geographic patterns of such connectivity-led agglomeration strategies taking the urban system in England and Wales as the testbed again. The hierarchical linkage clustering algorithm pairs city units with complementary infrastructural requirements where pairings mirror provision of inter-city connectivity links. We explored the robustness and persistence of these synthetic city-regions by performing a co-occurrence frequency analysis examining the recurrence of specific patterns across spatial scales. The findings provide evidence to believe a single-scale approach and focus when analyzing connectivity patterns without reasonable justifications, whether to inter- or intra-urban, only arbitrarily constrains the available solution space for a problem that would otherwise require a more concurrent consideration across spatial scales and distances. More importantly, when considering long-term ramifications of the size-cost balance allowing for potential productivity improvements due to improvements in connectivity, the results reinforce a need for a more comprehensive consideration of built-density in order to maintain overall size-cost balance.

8.2. Avenues for future research

In discussing potential avenues for future work, we outline two streams of works to be undertaken. One comprises works that address shortcomings of the current approach implemented in this volume to better refine the approach and findings. The other concerns works pertaining to exploration of new issues raised in this volume. We start with the former.

The studies presented in chapters 4 and 5 rely on estimating average aggregate responses of the overall urban systems, here those of England and Wales, Germany, and the Netherlands, at different spatial scales. Further work is required to evaluate and quantify the potential structural devision of the urban networks from a single coherent system to potentially a few geographically distinct regions of significantly different baseline productivities and dynamics. This work would entail the use of geographically weighted regression models for the estimation of regional scaling elasticities and ideal baseline human productions in parallel with varying spatial scales.

Similarly, the clustering exercise presented in chapter 6 currently considers mostly theoretical formulation of connectivity-based agglomerated city-regions. A more in-depth exploration of city-region development with considerably more nuanced clustering criteria is required to transform the approach from one designed to study the effects of spatial scales on geographic patterns of agglomeration to a flexible tool integrable with existing micro-scale planning models. These efforts would have to include further refinements of the placement of city-units, which is currently based on the centroid of unit polygon, to implement a more realistic city-unit center.

Finally, the current analysis has been performed based on total urbanized area taken as a proxy for the networked area of the mobility infrastructure. A more detailed examination of the overall road and transport network could enable a more direct quantification of inhabitants' average connectivity path and/or the lack of adequate connectivity through a direct examination of its geospatial network. This would provide a comparison for the quantification of connectivity provisions as the examination of the urbanized area scaling exponent might overestimate if the road/mobility network embedded within it is not adequately connected or financially accessible.

Moving on to the stream addressing new questions raised by the insights offered in this work, we begin with the observations regarding the coupling between scaling deviations of economic output and urbanized area. While the data from the Dutch, German, and English urban systems suggest urban units with a higher density than is expected of their size economically outperform their theoretical scaling expectation, further insights are needed as to the underlying dynamics governing this systemic coupling in the scaling deviations of economic output and urbanized area. Such attempts can begin by expanding the allometric comparison outlined in chapter 5 to urban systems that share fewer similarities with the European urban systems investigated in this work.

The core theme of this work has been examining the effects of spatial scales on the observation of power-law scaling exponents for which we have made use of one amongst many available scaling models. Even though we have pointed out in chapter 7 that the various scaling models that exist arrive at these scalings through differently framed formulations of the same ideas vis-à-vis connectivity and interactions, our results could ultimately remain limited by and specific to the choice of model. Future works on the effects of scale on these scaling dynamics need to also investigate the extent of agreement between these family of models in order to better isolate and understand the extent of the influence of spatial scales.

Lastly, a related area that was peripheral to the theme of this work and was hence left unexplored involves the real-life implementations of and quantification of *better connectivity*. We have only considered mobility and transport infrastructure abstracted through a notion of connectivity that governs the scaling dynamics. There remains questions on the effectiveness of various transport modes and other non-hardware mobility incentives in providing and contributing to this abstracted notion of connectivity. Further research may explore these through an investigation of the scaling deviations of city units controlling for transport infrastructure spending and modes within units.

A

Densification: An Energy-Based Digression

Material prepared for this chapter have been used in the following–

Arbabi, H., Mayfield, M. (2016). Urban and Rural—Population and Energy Consumption Dynamics in Local Authorities within England and Wales. *Buildings*, 6(3), 34.

Part of the overall utility of Bettencourt's particular model formulation rests in its bypass of direct measurements of W, and hence $Y - W$. In realty, empirical measurement or estimation of mobility-related W across spatial scales is difficult. This in fact appears to be the case for comprehensive measurements of all energy consumption across spatial scales. In the particular case of England and Wales, this is partly due to the unwillingness of energy providers to share or collect high resolution data and the privacy and accuracy concerns when estimating consumption at high spatial resolutions, e.g. individual gas or electricity meters.[1] For these reasons, we have to content ourselves with the data available at the LAU1 units as the multi-scale examination of scaling patterns for energy consumption, whether it be transport-related or not, is not as conveniently feasible.

[1] At the time, the lowest level at which the Department for Business, Energy & Industrial Strategy and ONS provide full aggregated energy consumption with reliable confidence, by fuel and consumption type, is LAU1. Although, modeled domestic consumption is available at MSOA boundaries and single-year experimental data pertaining to individual meters aggregated at a postcode level (Gregory, 2014).

In chapter 5 and 6, we highlighted the observation that taking up an agglomeration-based perspective would suggest economic improvements in under-performing parts of the urban network in England and Wales would have to be accompanied by plans to compact and densify the urbanized areas in order to avoid an economic improvement at expense of higher energetic mobility tolls, W. In this digressional appendix, we review a parallel case for urban density, and specifically the spatial homogeneity of increased population density, from an energetic perspective. This is presented as an examination of the presence of similar power-law relationships pertaining to energy consumption and the effects of density at the only compatible boundary definition, i.e. local administrative units level 1 (LAU1), where data is available.

A.1. Urban or rural

So far, we have been using the method developed by Clauset (2009) to filter out small units. In this appendix, in addition to the use of the population cut-off value, we also estimate Gini coefficient for population density in each LAU1 to delineate urban from rural.[2] The reasons are twofold. Firstly, since it is the effects of density and densification we are interested in, the Gini provides a simple indicator. Secondly, we use the indicator because the transport consumption data have been estimated over the overall extent of each local authority, A, rather than our previously used urbanized area indicator, A_n. As such, for larger, more *rural*, and hence sparser LAU1 units, transport consumption would be more sensitive to the distance separating individual patches of urbanized area. We account for these by calculating the Gini coefficient for each LAU1 unit based on the overall population density, $\frac{N}{A}$ rather than $\frac{N}{A_n}$, of all MSOAs building

[2] Gini coefficient, an index of (in)equality, provides a measure of heterogeneity within a distribution, see Equation A.1.

up that local authority according to

$$Gini_{LAU1} = \frac{\sum_i^n \sum_j^n |(\frac{N}{A})_i - (\frac{N}{A})_j|}{2n^2(\bar{\frac{N}{A}})} \tag{A.1}$$

where n and $(\frac{N}{A})$ are the number and population density of MSOAs respectively and $(\bar{\frac{N}{A}})$ denotes average population density of all comprising MSOAs for each LAU1 unit (Schwarz, 2010; Tsai, 2005). The index provides a measure of inequality of overall population density across the MSOAs of each authority unit. Values closer to 0 are indicative of a more uniform distribution of population over the area of the LAU1 whereas those closer to unity point to an extreme disparity between the population density of different MSOAs in the same LAU1. As such, it is expected that rural authorities that have only a handful of MSOAs where the settlements are located with larger densities and numerous MSOAs enveloping unpopulated and empty land would exhibit larger Gini coefficients compared with their urban counterparts which are not expected to be punctuated by sparsely populated MSOAs. Consequently, the choice of the Gini coefficient based on population density of constituting MSOAs enables an assessment of the dispersion and homogeneity of population distribution within each LAU1. For the urban/rural threshold presented here, a cutoff value of 0.4 has been utilized with those below marked as urban and the rest rural. It should be noted that the method utilized here matches that of the Department for Environment, Food & Rural Affairs, for a majority of the local authorities identifying 160 urban LAU1s as opposed to the 181 within the first three urban tiers of their classification. Figure A.1 shows the visual implementation of the Gini coefficient over the LAU1s in England and Wales.[3]

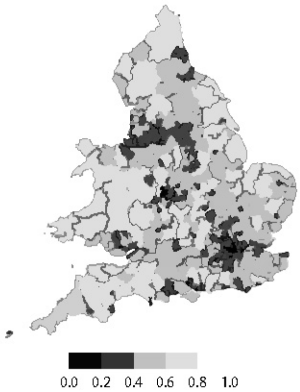


Figure A.1: LAU1 units across England and Wales color mapped based on their respective Gini coefficients – Contains National Statistics and OS data © Crown copyright and database right 2016.

[3] For a map of the 2011 urban-rural classification see Department for Environment, Food & Rural Affairs (2015)

A.2. Energy consumption scaling

Table A.1 summarizes the result of a similar OLS analysis performed for the log-transformed total fuel consumption (TF), total transport consumption (TT),[4] total industrial and commercial consumption (TIC), and total domestic consumption (TD) including its two major components, i.e. domestic gas (TDG) and electricity (TDE), against population for LAU1 units in the year 2011.

[4] This is mostly an aggregation of modeled road transport consumption and rail, see (Bircknell, 2018).

The consumption across sectors, as expected, corresponds closely with the total resident population with statistically significant fits.[5] The scaling regime, that is the exponent of the power-law, also appears to be broadly similar across consumption types when considering all LAU1 units or those with populations exceeding the cut-off

[5] The goodness of fits, however, is not as tight as those observed for economic output and urbanized area previously.

Table A.1: Regression analysis report for log-log relationship between energy consumptions and population for LAU1 units.

	TF	TT	TIC	TD	TDG	TDE
Sample Size	346 (all LAU1s)					
$\boldsymbol{\beta}$	0.83	0.68	0.92	0.94	1.05	0.90
$\boldsymbol{R^2}$	0.63	0.36	0.44	0.96	0.88	0.95
95%CI	[0.77, 0.90]	[0.58, 0.77]	[0.81, 1.03]	[0.92, 0.96]	[1.01, 1.10]	[0.88, 0.92]
Sample Size	215 (as per Table 4.1)					
$\boldsymbol{\beta}$	0.88	0.77	0.98	0.97	1.03	0.94
$\boldsymbol{R^2}$	0.63	0.49	0.42	0.97	0.90	0.96
95%CI	[0.79, 0.98]	[0.67, 0.88]	[0.83, 1.14]	[0.95, 1.0]	[0.99, 1.08]	[0.91, 0.96]
Sample Size	160 ($Gini \leq 0.4$)					
$\boldsymbol{\beta}$	0.91	0.83	0.98	0.93	0.94	0.93
$\boldsymbol{R^2}$	0.81	0.55	0.6	0.95	0.94	0.97
95%CI	[0.84, 0.98]	[0.72, 0.95]	[0.85, 1.10]	[0.90, 0.96]	[0.91, 0.98]	[0.90, 0.95]
Sample Size	186 ($Gini > 0.4$)					
$\boldsymbol{\beta}$	0.93	0.82	1.01	0.98	1.09	0.93
$\boldsymbol{R^2}$	0.59	0.48	0.36	0.97	0.83	0.95
95%CI	[0.82, 1.04]	[0.70, 0.95]	[0.81, 1.20]	[0.95, 1.00]	[1.02, 1.16]	[0.90, 0.96]

with the Greater London Authority aggregated. From our tabulated results, for the exception of domestic gas, the energy consumption follows a sub-linear to linear regime.[6] The TDG consumption, however, within its 95%CI exhibits a linear to super-linear behavior especially for units with $Gini_{LAU1} > 0.4$ indicating increases in *rural* population is more likely to result in proportional or rising increases in domestic gas consumption. Moreover, although the exponents estimated for the transport-related consumption do seem to agree for both rural and urban regions in a sub-linear scaling, it can be seen from Figure A.2 that baseline prevalence of road consumption in general is higher within the rural LAU1 units.

[6] It is important to note that although the SRM develops a super-linear scaling of energy *dissipated* over the infrastructure, the data available and presented only accounts for individual consumption/demand and not the overall dissipated over the network.

A.2.1. Consumption intensity and population density

Recalling the general theoretical formulation of the scaling laws

$$F(N) = F_0 N^{\beta}, \tag{A.2}$$

the scaling for urban energy consumption and area can be combined and rearranged for a general description of variations of energy intensity against population density in

$$\frac{E}{N} = \rho(\frac{N}{A})^{\frac{\beta_E - 1}{1-\beta_A}} \tag{A.3}$$

Urban and rural Energy consumption patterns

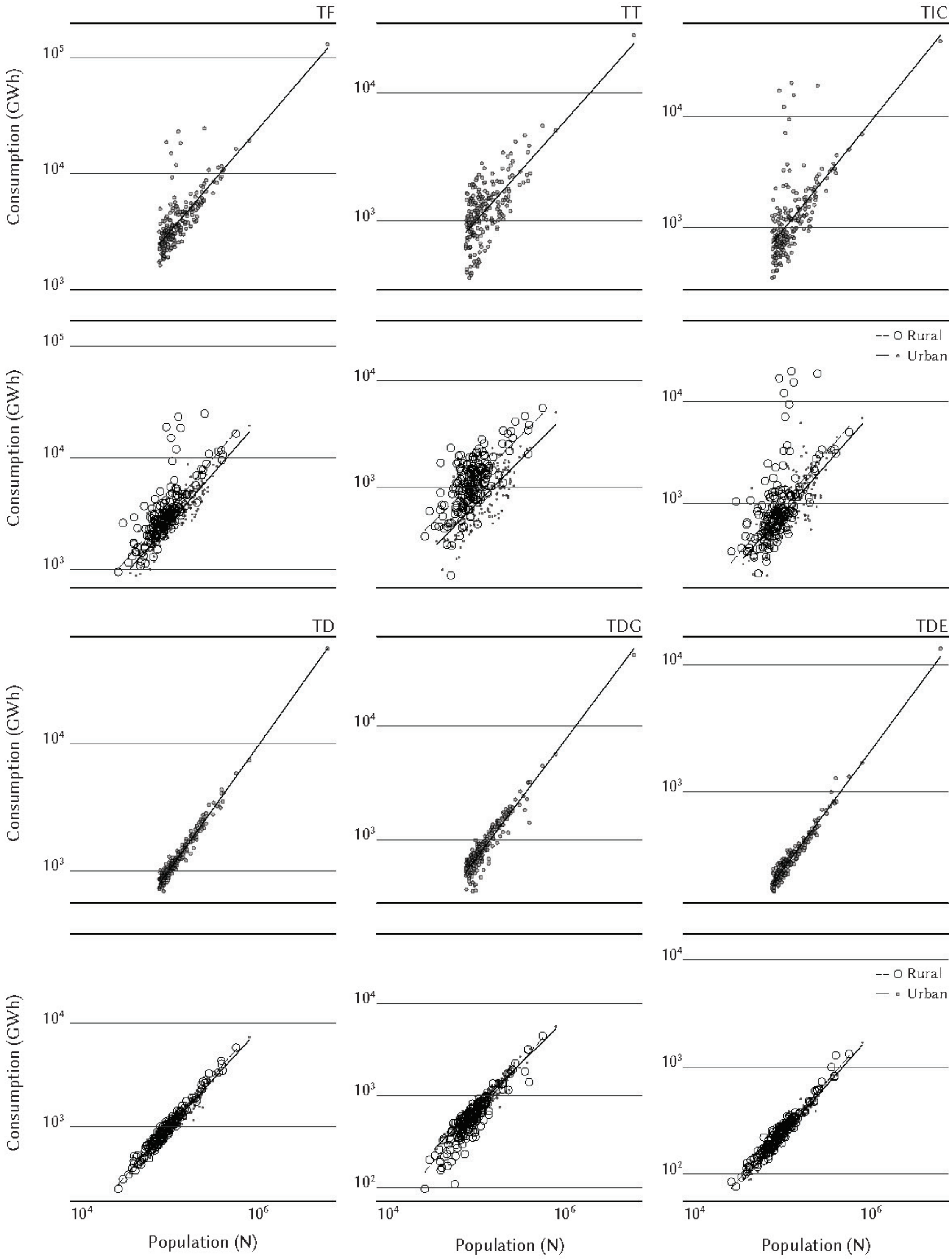


Figure A.2: Log-log variations of energy consumption (GWh) against population – for each consumption category top panels correspond with LAU1s as per table 4.1 while urban and rural LAU1s have respectively been denoted by solid and hollow data points in the bottom panels.

where E and A represent energy consumption and city overall area corresponding to population N, ρ an aggregate pre-factor incorporating E_0 and A_0, the baseline prevalences of energy consumption and city area from Equation A.2, and β_E and β_A the scaling exponent for each indicator. As can be seen from Equation A.3, to obtain inverse power-law correlations[7] where $\frac{\beta_E - 1}{1 - \beta_A} < 0$, similar to that observed by Newman and Kenworthy (1989, 1999) for private transport consumption two decades ago, both energy consumption and area of the city units need to display a sub-linear scaling with population, i.e. $0 < \beta_E, \beta_A < 1$.

[7] The expectation of an inverse relationship between per capita consumption and density is fundamental to the belief that bigger and denser cities are more efficient.

Therefore, the dissimilarity of scaling patterns for the TDG consumption in urban and rural authorities discussed previously would also be present in the per capita variations of consumption across the three sectors with population density. As can be seen from Figure A.3, per capita consumptions of TT and TDE do display a similar decreasing trend with population density to that observed by Newman and Kenworthy (1989, 1999) where more uniformly dense urban areas constitute the stable low-consuming long tail as the rural LAU1s populate the energy intensive area to the left. This is while the per capita consumption TDG experiences an initial rise with the increasing population density before starting on an incredibly slow decreasing pattern with the shift from rural LAU1s to urban ones. This is consistent with the behavior expected based on the exponent values estimated for them previously and those we have seen regarding the scaling of A_n, despite having used A in calculations here. It should, however, be noted that although the urban/rural classification does account for and explain the split between the increasing/decreasing response, based on the power curves fitted here, changes in population density seem to explain less than half the variations in per capita consumption. It can also be gleaned from the data that the total consumption across LAU1s remains more or less proportional. As such, rural authorities housing about 45% of the population are also responsible for roughly 44% of the overall domestic consumption. This is in spite of the differences and trends in the domestic per capita consumption of urban/rural authorities and is mostly a direct result of the overarching population scaling effects, see Figure A.2. The same, however, cannot be said of the transport consumption where, despite their population share, the units with $Gini_{LAU1} > 0.4$ constitute 60% of the consumption.

Per capita consumption and population density

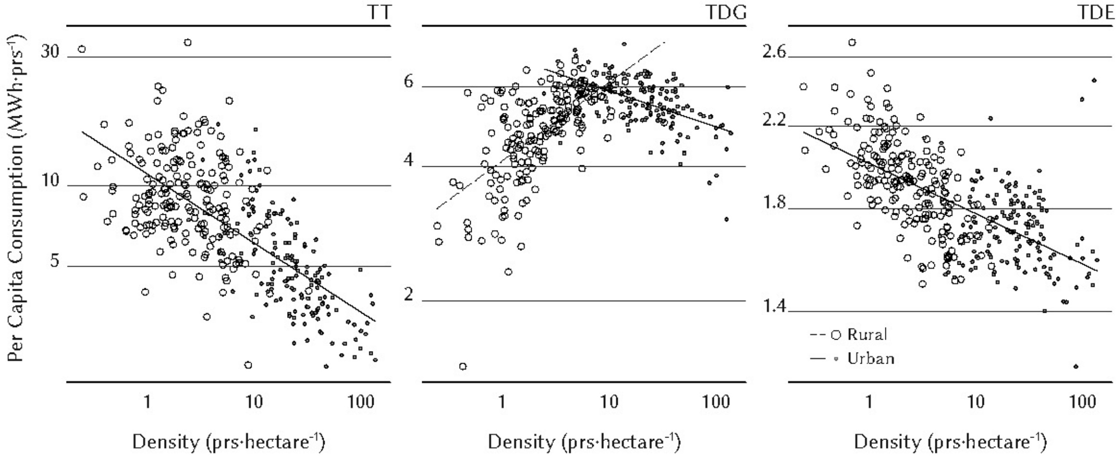


Figure A.3: Per capita variations of TT, TDG, and TDE with population density on logarithmic axes – urban and rural LAU1s have been denoted by solid and hollow data points, respectively.

A.3. Densification from an energy-based perspective

The first issue that requires further explanation is the quality of the regression fits for the TT consumption. Unlike those for TDG and TDE, the R^2 of the fitted lines for transport consumption against population are weak with population explaining only around 50% of the variance in transport consumption. This is also the case for the per capita consumption in the sector versus population density. Overall scaling effects are, however, undoubtedly present albeit far noisier than those seen for other consumption types. The transport-related consumption can be argued to depend on several competing factors ranging from fuel prices to vehicle ownership, road and public transport network quality, and potential traffic waiting hours (Gordon, 2008). The low R^2 values and the scatter clouds present in the per capita consumption and density plots for the other two sectors, however, are most likely due to major departures from the expected scaling relations for land area, A, as compared with, say, urbanized area, A_n.

Secondly, the super-linear exponent observed for the domestic gas consumption in rural, read less homogeneously dense, regions can perhaps be explained by looking at the end-use of gas in the UK. The consumption of domestic gas has been indicated to largely address the space heating demand which constitutes about 70% of total domestic demand (Palmer and Cooper, 2011, 2012). From a thermodynamic perspective, space heating, unlike other domestic demands that are more individual-driven, would only enjoy the effects of economies of scale when subject to more compact constructions (Hui, 2001; W. O'Brien, Kennedy, Athienitis, and Kesik, 2010) which usually implies a smaller surface area thermodynamically and the more effective implementation of efficient energy networks. In an urban context, further increases in population can be taken as an indicator for increasing compactness of the built form and therefore higher consumption efficiencies, hence the slightly sub-linear response of the domestic gas in the uniformly dense LAU1s. In a sparse rural setting, however, increases in total population do not necessarily translate into more compact morphologies given the nature of such settlements. In fact, a visual comparison of Gini coefficient and aggregate population density of the LAU1s shows how the higher density urban cities also have lower Gini coefficients meaning the entirety of the population in them is focused around fewer central cores as opposed to the larger coefficients calculated for the rural authorities indicating the existence of separate and in some instances highly dispersed dense settlements, Figure A.4. Consequently, in the absence of a decrease in surface-to-volume ratios, i.e. densification and compaction, increases in population would not be accompanied by the expected sub-linear scaling indicative of increasing efficiencies.

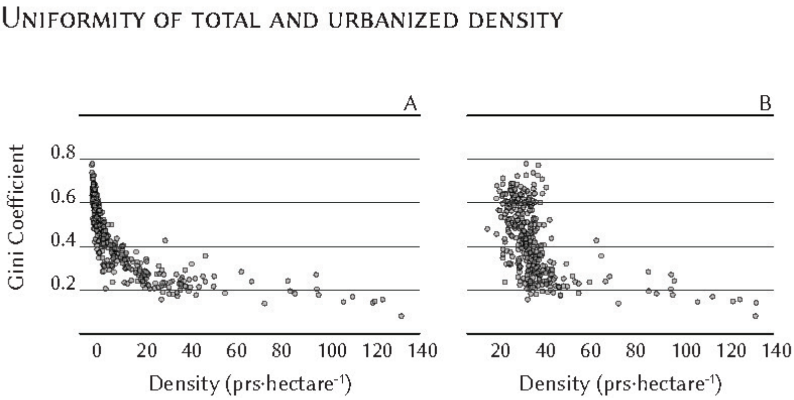


Figure A.4: Plot of Gini coefficient against total density (A) and urbanized area density (B).

What we see common across all consumption categories, and visualized for TT, TDE, and TDG in Figure A.3, is the observation that although there might not be significant savings to be had, from an energy consumption perspective, in further densifying already uniformly-dense units, the agglomeration-supported savings are found in the densification of heterogeneous population scattered over sprawling settlements. This is also consistent with our mobility- and output-driven line of inquiry in chapters 4 and 5.

B

Caveats and Qualifications: A Sensitivity Analysis

The adoption of Bettencourt's social reactor model (SRM) and our overall scaling point of view have enabled this multi-scale analysis with only a handful of variables, namely, gross value-added (GVA), urbanized area, and population. In this appendix, we review potential sources of sensitivity and uncertainty in the analysis some of which, such as the simplicity of the GVA approximations, have already been alluded to in preceding chapters.

The caveats and qualifications to be made in this appendix can be broadly grouped into the following categories:

- those concerned with the quality of the input data used, referring both to data imported unchanged from external sources and those altered and scaled across different scales,[1] and the sensitivity of the subsequent findings to it,
- those concerned with the definition of city units and the way choice of city bounding box and eligible population might influence and alter conclusions, and finally
- those addressing underlying assumptions made throughout as a foundation of the modeling paradigm and implicitly incorporated through our particular interpretations of them.

[1] This almost exclusively refers to the area- and population-based breakdown of the GVA values from the NUTS3 boundaries and aggregated at other boundaries as described in chapter 3.

The appendix will outline these in the same order set out above. We will first briefly review our input parameters and the potential effect their variation might have on our analysis. This we follow with a sensitivity analysis set up as to quantify some of this influence. We then address the ramifications of a change in the specific boundaries[2] used for city units by looking at a case-study of the Local Authority units in the six combined city-regions of the Northern Powerhouse. Finally, we conclude with a discussion of the generality of the findings presented in the preceding chapters as a function of the subjective interpretation of some of social reactor model (SRM)'s assumptions and model outputs.

[2] These changes refer to those affecting individual city units territorially rather than the overall spatial scale at which cities have been defined for a given boundary definition. Although, similar to the first category of caveats and qualifications, these result in variations in cities' population, economic output, and urbanized area, unlike the first category these variations are not due to calculation inaccuracies or uncertainties.

As outlined in chapter 3, the primary input variables for the scaling analysis underpinning this work are those quantifying a city's population, N, its economic output, Y, and its urbanized area, A_n. Providing for these parameters, we have used a breakdown of 2011 Census population at 1×1 km^2 grid for estimating population at different scales, the sub-national GVA estimates for NUTS3 areas as the base for economic output estimations, and the CORINE land cover dataset for the quantification of urbanized area. Of these three, urbanized area is the most reliably invariable. The satellite based CORINE data and its methodology are fairly mature (European Environment Agency, 2000) and although detection and allocation of urbanized land might contain errors these remain external to our approach.

For population, the choice of the grid as the common base to estimate unit population across scales minimizes internal calculation inaccuracies as estimations for the density-based units relies on a simple addition of values across conjoined grid cells. It is only the administrative and functional boundaries the population estimations for which require dealing with the fractured grid cells intersected by units boundaries. However, the significantly smaller number of such fractured boundary cells compared with those remaining within units intact means that such population variations are negligible as a fraction of each unit's total population.

In contrast to these, because the GVA is only recorded at the NUTS3 boundary,[3] to enable the deployment of the model across scales, values have had to be scaled down to higher-resolution spatial units,i.e. the population gird cells, enabling estimation of economic output for units of differing spatial scales from a common base. In absence of extensive complementary data and following the OECD's GIS-based approach, this has been undertaken using linear proportionality of intersected areas and populations.[4] As mentioned in chapter 3 and pointed out by Smith (2014) such methods while convenient can cause miscalculations at smaller units.

[3] This has changed since the conclusion of this study. Office for National Statistics (ONS) has since released experimental residence-based estimations of gross value-added for LAU1 units. Although still requiring the same break-down approach to use with the density-based units, the overall smaller nature of LAU1 units would potentially minimize density-related inaccuracies and limit them to a smaller number of units.

[4] Refer back to chapter 3, Equation 3.13.

B.1. Exploring the sensitivity of results to economic output

The question that arises when reviewing the calculation inaccuracies in the input parameters is then whether these have the potential to alter our estimations of

- the overall scaling exponent and hence the scaling regime associated with the parameter, and
- the approximated G^* and hence the overall infrastructural needs of the urban network.

B.1.1. Sensitivity analysis design

As mentioned in the previous section, of the three urban variables involved, city units economic output, Y, is the most likely to have been contaminated by calculation and scaling errors. As such, in order to gain some understanding of the importance and magnitude of these, we follow a bootstrapping process whereby we subject this input urban metric to random variations re-estimating values for scaling exponents and optimal G^* over a large number of simulations. Since the overall magnitude and systemic distribution of these variations are unknown, a conservative approach would see the scaling of each city's economic output according to

$$Y_i^{'} = (1 + \omega)Y_i \tag{B.1}$$

where $Y_i^{'}$ denotes the adjusted economic output of city i, Y_i its original estimated output as outlined in chapter 3, and ω a magnitude multiplier drawn randomly from a distribution with an average of zero.[5] Figure B.1 shows three separate distributions used when drawing ω. This approach and the choice of distributions mean that each city, regardless of its size, could either be completely discarded ($\omega \leq -1$) or see a doubling of its economic output in each iteration.[6] However, knowledge of the potential sources for variations in the estimates of Y dictate that these are more likely to affect smaller city units as these would inevitably be comprised of fewer grid cells meaning a more prominent influence of the sparse population densities and linear scaling uniform density assumptions. To account for this, we consider an alternative formulation of Equation B.1

$$Y_i^{'} = (1 + \omega P(N \geq x))Y_i \tag{B.2}$$

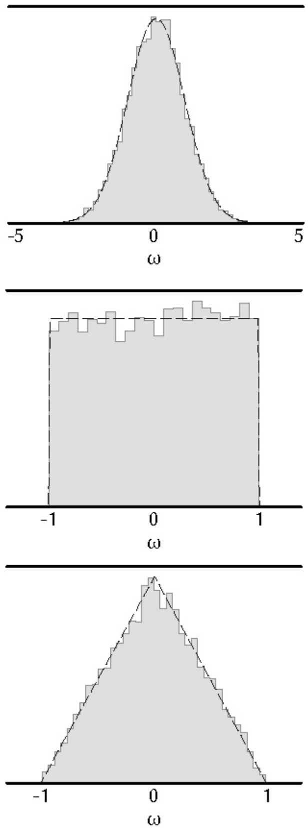


Figure B.1: Distributions used for ω, from top to bottom, standard Gaussian, standard uniform, and triangular.

[5] Note that if ω is drawn from a *zero-centered* distribution, and applied in Equation B.1, one would not expect a significantly different average for the distribution of estimated $\beta_{Y'}$ over many simulation runs from the original β_Y and the formulation simply helps determining the standard deviation of $\beta_{Y'}$ distribution.

[6] Strictly speaking, it is only the standard uniform and triangular distribution that are bounded by $[-1, 1)$ domain.

where the additional $P(N \geq x)$ denotes the complementary cumulative probability for a city of population N effectively scaling the added noise ω so as to attenuate the variations for cities of larger populations. Figure B.2 shows the difference in the variation regimes under Equations B.1 and B.2 where the shaded area represents the available variation domain for an ω drawn from the uniform distribution.

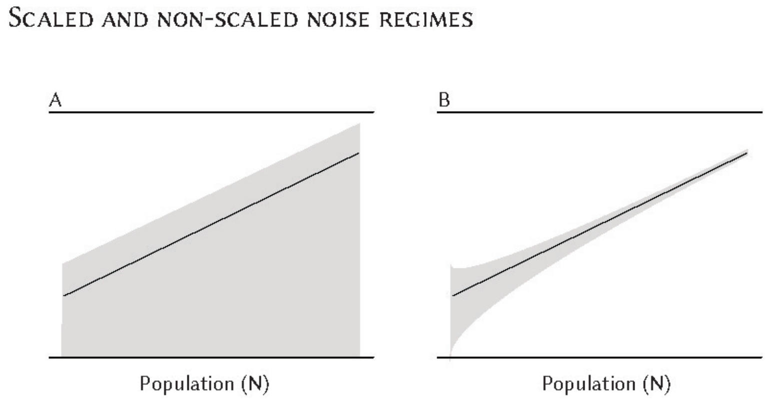


Figure B.2: Schematic showing the variation regimes allowed for economic output as a function of population where ω is drawn from the uniform distribution for (A) non-scaled noise and (B) population-scaled noise – note that the axes are in arbitrary units and logarithmic.

It should be noted that, although the sources of variation mentioned are common to the estimates of both output and population, we only consider variations in Y rather than simultaneously perturbing both city population and output since estimation of one, i.e. Y, is already tied to the other because of the $1 \times 1\text{km}^2$ grid and the subtle variations of N would already have been included in letting Y values fluctuate. Also, estimations of urbanized area, A_n, are satellite-based and unlike output estimates are independent of city population in their derivation. We therefore do not have a rational basis to assume and/or introduce arbitrary variations to city estimates and an exploration of estimation errors in them falls beyond the scope of the present work.

B.1.2. Average urban network response

We begin by looking at the effect of output variations on the OLS estimates of $\beta_{Y'}$.[7] Figure B.3 shows the average estimated exponent and the standard deviation of its distribution over multiple simulation runs. As can be seen, the average estimates pertaining to the non-scaled variations from all three countries do coincide fairly closely with those reviewed previously in chapter 5. The population-scaled estimates, on the other hand, show an average $\beta_{Y'}$ larger than β_Y with the estimates for density-based boundary definitions in Ger-

[7] Note that values presented pertaining to this sensitivity analysis are the aggregated results of 10,000 simulation runs using the same units as in chapter 5. Due to the very small sample size, OECD units in the the Netherlands have not been considered.

many now even closer to the idealized theoretical expectation. This is for the exception of the estimates for URBAUD and OECD boundaries where the average $\beta_{Y'}$ is slightly smaller than the originally estimated exponents previously.

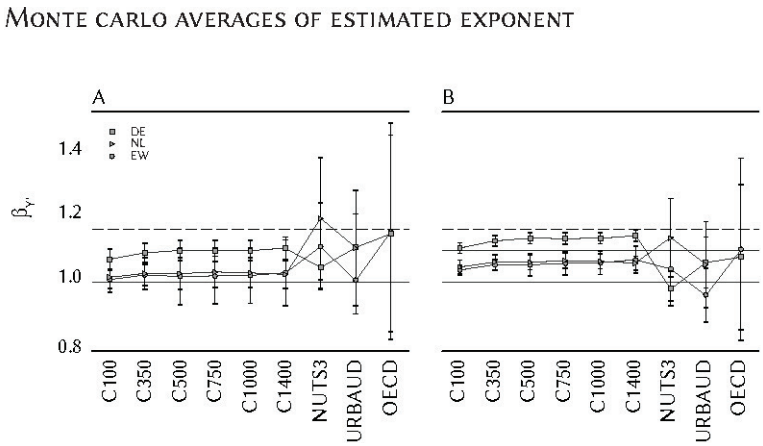


Figure B.3: Point plots showing the average estimated $\beta_{Y'}$ and its standard deviation for each boundary using (A) non-scaled variations and (B) size-attenuated variations with dashed lines indicating theoretically ideal threshold for $D = 2$ and $H = 1$ – note that the values are averaged over results for all three distributions of ω.

Another, perhaps more subtle, difference between the estimates with and without size-attenuation is the larger standard deviations of the estimates for England and Wales in the non-scaled simulations. These are due to the possibility in the non-scaled scenario of excluding the London area. We have repeatedly noted the perceived uniqueness of London in England and Wales's urban network. Unlike in Germany and the the Netherlands where there is not a glaring population divide between the two most populated units under each boundary definition, a significant chasm exists between the population size of London and the other units in England and Wales.[8] As such, exclusion of London from the regression, unlike its German or Dutch counterparts, results in the loss of a leading point from the data with a noticeably larger effect on the OLS estimation of the exponent which are captured in the larger standard deviation seen in inset A of Figure B.3.

[8] Using URBAUD units as an illustration, the most populated units in Germany and the the Netherlands are 1 and almost 2 times more populous than their second cities, respectively. Meanwhile, London is over four times the second largest URBAUD unit in England and Wales.

What is of more interest is that, despite the variations in output and resulting changes in the estimated exponent, the three countries primarily maintain their relative order with Germany still exhibiting exponents larger than the other two. The England and Wales population-scaled estimates in Figure B.3 might suggest a better mobility and accessibility provisions were we to assume significant mis-allocation of economic output in smaller units. Even in the event of such mis-allocation, the English and Welsh urban network remains behind when compared with that of Germany.

B.1.3. Theoretical ideal and the distribution of infrastructural needs

Next to consider are potential changes to the fixed-slope estimations of Y_0 as a result of the random variations and hence the fluctuations of the estimated G^* and cities' η. Overall smaller average values for the idealized $G^*_{Y'}$ compared with the estimates used previously would not be surprising. As we have seen, the average estimates for $\beta_{Y'}$ are slightly larger than the original β_Y. This also implies that the estimates of the intercept, $Y_0^{'}$, are likely to be smaller – note that given the mean of the data remains more or less stationary, an increase in slope requires a decrease in the intercept. Same decrease in the estimated values would also be expected for the intercept of the fixed-slope regressions required for estimating $G^*_{Y'}$, see Equation 4.2. Given that we do not consider changes in urbanized area, this reduction in the intercept value would directly result in the smaller estimates for the baseline productivity, Figure B.4.

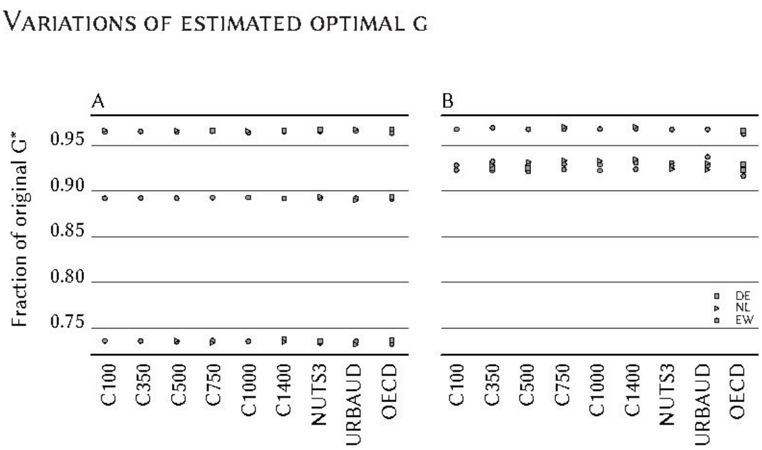


Figure B.4: Plots showing the fraction $\frac{G^*_{Y'}}{G^*}$ for (A) non-scaled variations and (B) size-attenuated variations.

These smaller estimates mean that a larger number of units could potentially have been seen as requiring a compaction of built-area[9] than previously suggested in chapters 4 and 5. To gain a feel for the number of units affected, we look at the percentage of units for which the model recommendation would change if we were to use the average $G^*_{Y'}$s from the simulations rather than the originally estimated G^*.[10] Figures B.5, B.6, and B.7 show these percentages for the different distributions of ω. When considering the size-attenuated estimates, regardless of the distribution used for ω, the model recommendation for roughly 10% of units would change from requiring better mobility to more compact urbanized extents. While these ratios are more or less stable for the size-attenuated simulations, the non-scaled results vary from 5% to 40% depending on the choice of the distribution used. This means that the previous observation

[9] Alternatively, mitigations of larger-than-ideal Gs could be achieved through increased energy efficiency of the transport modes delivering mobility in such units.

[10] Note that for these we are only changing the perceived point of optimum with the original estimations of economic output, urbanized area, and population.

Percentage of units with altered infrastructural need

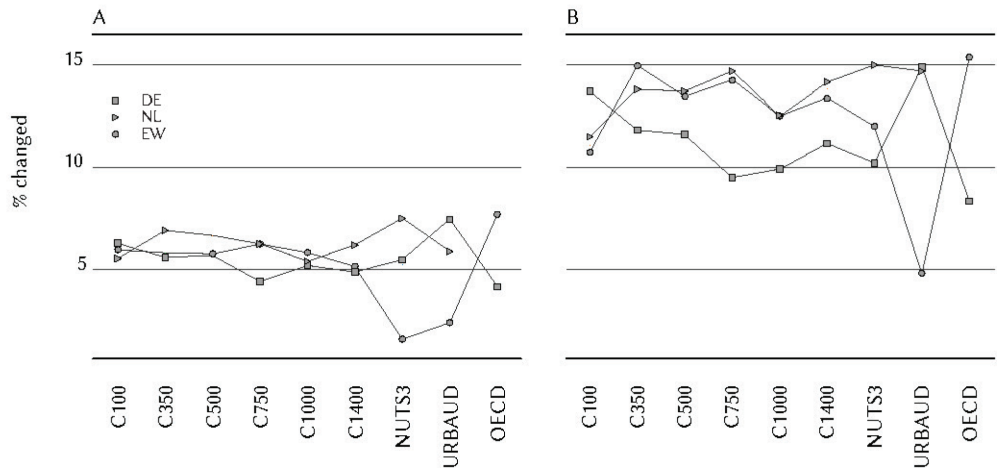


Figure B.5: Point plots showing the percentage of units where the model recommendation changes due to a change in G^* estimation using (A) non-scaled variations and (B) size-attenuated variations – ω from a standard Gaussian distribution.

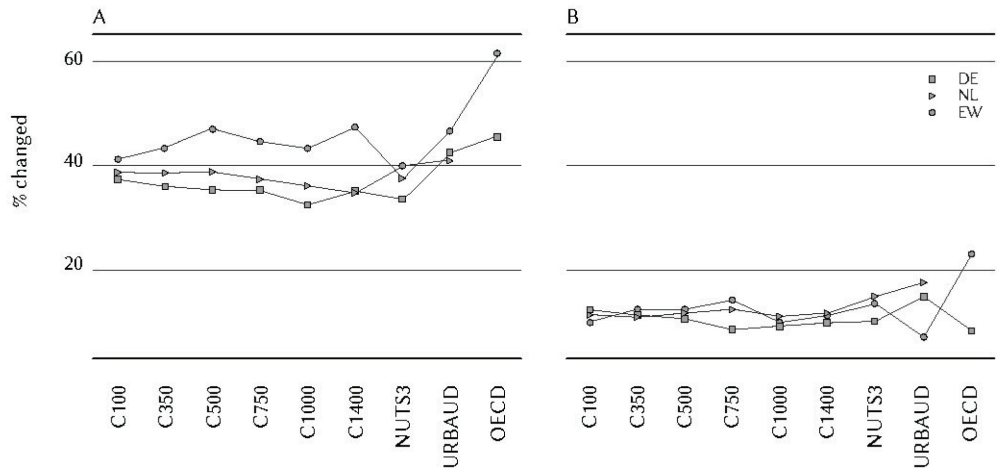


Figure B.6: Point plots showing the percentage of units where the model recommendation changes due to a change in G^* estimation using (A) non-scaled variations and (B) size-attenuated variations – ω from a uniform distribution.

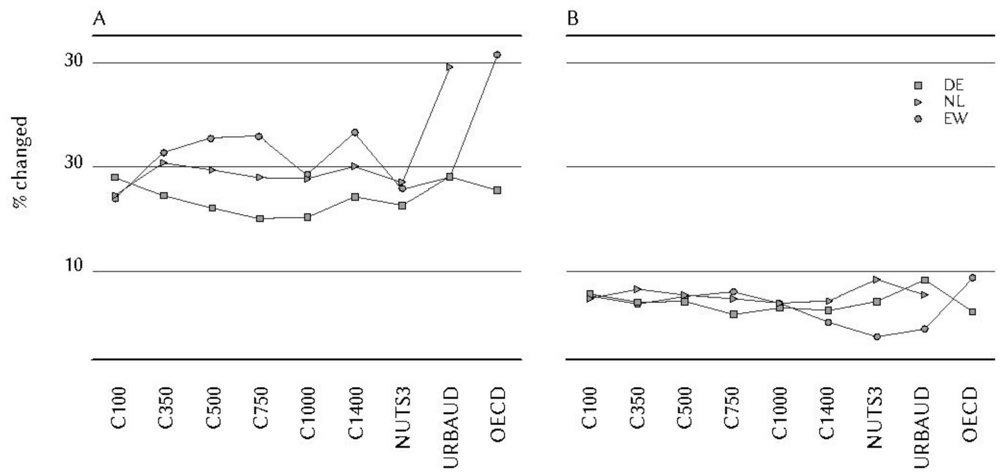


Figure B.7: Point plots showing the percentage of units where the model recommendation changes due to a change in G^* estimation using (A) non-scaled variations and (B) size-attenuated variations – ω from a triangular distribution.

in chapters 4 and 5 regarding the majority of units in England and Wales exhibiting a lack of adequate mobility and hence requiring better provisions of mobility and access infrastructure might require further qualifications. A very significant number of units may appear to be affected by the choice of the point of optimality. What is crucial to keep in mind, however, is the more fundamental observation that the urban network in England and Wales remains systematically and on average less homogeneously mixing and accessible than would be required to deliver desired economic agglomeration effects as reflected in the values of $\beta_{Y'}$ already reviewed.

B.2. Exploring the sensitivity of results to city boundary and population

Another difficulty alluded to from the outset is the identification of the appropriate boundary for each city unit and hence the estimations and allocation of proper infrastructure network area and population. To address this issue and provide a comparison for potential variations of the results, we consider three simple scenarios determining the boundary, or rather the bounding box, for the proposed city-regions in the Northern Powerhouse:

i. a purely administrative one following the LAU1s boundaries and their aggregated combined authorities and city-regions, Figure B.8,
ii. one based on the ratio of the built-up area to the total area[11] within the lower-layer super output areas (LSOAs) in each region, and finally
iii. another based on the population density within the LSOAs in each unit.

[11] Note that for the purposes of this appendix, *area* refers to the total areal extent of each unit and not the urbanized areas used previously.

Administrative boundaries. Under this scenario, each of the six city-regions of the Northern Powerhouse is assumed to be the city bounded by the union of administrative boundaries of their comprising LAU1s. Figure B.9 shows the administrative boundaries of the 36 northern LAU1s and the city-regions they form. As can be seen, the administrative boundaries do not necessarily correspond to the boundaries of the inhabited and economically active city. This is best illustrated considering the local authority of Sheffield where the western portion of the area within the administrative boundary is part of the Peak District National Park.

Similar effects are also observed in inherently rural parts where the administrative boundaries are not that of a coherent urban settle-

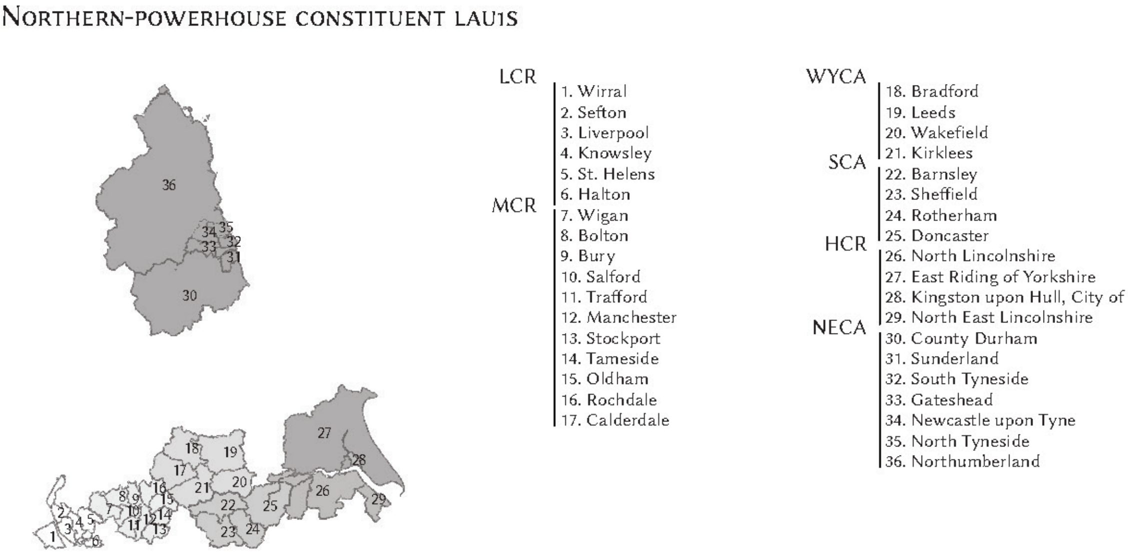


Figure B.8: Map of the Northern Powerhouse (NPH) and the local administrative units level 1 units comprising them – contains OS data © Crown copyright 2016.

ment or metropolitan area but rather a collective of sparsely scattered rural settlements, e.g. the local authorities of Northumberland and East Riding of Yorkshire. Using overall area of the administrative boundaries in these cases would then be in direct violation of one of Bettencourt's theoretical model assumptions, i.e. homogeneously mixing population, and in essence over-estimate the empirical value of G in these areas. We intentionally use these areas here to explore whether the model's interpretations can still be rationalized even when using incompatible inputs with the model's internal assumptions.

Built-up cover ratio-based boundaries. For this scenario, the total area in each local authority is calculated as the sum of the areas of all the LSOAs the ratio of the built-up area which is not less than a standard deviation below the average effectively excluding the rural portion of the population and land area. The GVA estimates for each LAU1 and city-region, however, is kept the same as those utilized for scenario i. The decision to limit area and population that is considered to be contributing to the performance of cities but not the share of the GVA is mostly based on the understandings that the social interactions and movements responsible for the production of the GVA are still those taking place within the city proper and not the hinterlands where perhaps part of the labor might take place. This is compatible with the common assumption that the production of the traded goods takes place within the central business district for urban areas (Henderson, 1974).

Population density-based boundaries. In this scenario we use the same procedure as the previous scenario with the difference of

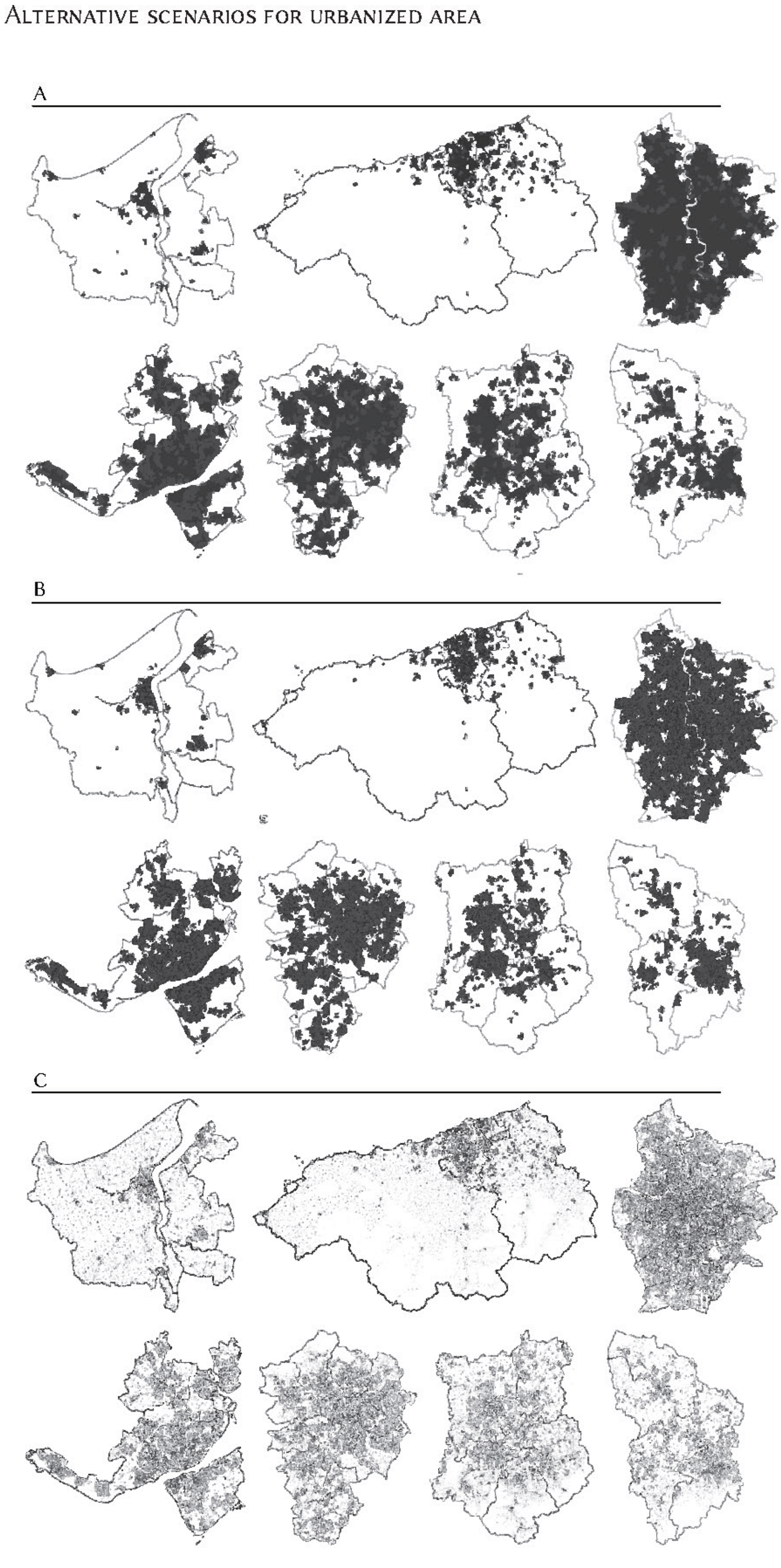


Figure B.9: Maps of the Northern Powerhouse (NPH) and the Greater London Authority for (A) scenario iii, (B) scenario ii, each with the extents considered for population and area highlighted, and (C) scenario i, with building footprints superimposed (see Figure B.8 for a correctly-scaled representation of the Northern Powerhouse) – contains OS data © Crown copyright 2016.

applying the standard deviation criteria to the population density of the LSOAs rather than the built-up cover ratio, Figure B.9[12].

[12] Note that since we are not repeating this exercise for all units in the England and Wales urban network, we assume London to act as a theoretically ideal unit such that $G^* = G_{EnglandandWales}$ for convenience.

B.2.1. Findings

Scenario i. Figure B.10 shows the estimated values of G for each of the six city-regions within the NPH and the Greater London Authority with the dashed line marking the threshold $G_{max} \approx 8G^*$ based on the G of the England and Wales. It is clear from the results that, under the first scenario, all six member regions of the NPH exhibit Gs higher than that of London with HCR and the NECA showing values higher than the threshold of positive cost-benefit balance within the city. Estimated values of G for the local authorities within each city-region, however, show a wider variation, Figure B.11, with a few, e.g. Salford, South Tynside, and North Tyneside, showing performance levels close to that of London.

Scenario ii. The change in the boundary considerations from one scenario to another provides a more nuanced picture of the performance balance among the regions of the NPH and England and Wales. Under the adjusted area, which in essence only accounts for certain more densely built-up LSOAs and their population, the balance of social benefits and energy dissipation costs across the NPH appears much closer to that of London which is more in tune with model's expectation and our previous observations that viable units would cluster close to the value of G^*[13]. The main difference, however, apart from the movement of HCR and NECA far below the maximum threshold, is the positioning of the estimated G for the LCR still close to that of London but now below it. One can also observe from Figure B.11 that the growing number of units with Gs below that of London are the de-facto centers of the city-regions to which they belong, i.e. Liverpool, Manchester, Sheffield, Kingston upon Hull, and Newcastle upon Tyne.

[13] Remember that in our previous examinations of England and Wales urban system London exhibited slightly more than ideal estimates of G.

Scenario iii. The estimates under the third and final scenario behave similar to that of the second. However, given the usually larger number of LSOAs included under this scenario, the population included in this scenario is higher than that of scenario ii for most LAU1s. Coupled with the higher sensitivity of estimated Gs to population than area ($G_{est} \propto A$ compared with $G_{est} \propto N^{-2}$), lower values of the estimated G and higher number of units with estimates below that of London are expected and more compatible with our prior estimation using the correct values of urbanized area. Scenario iii thus estimates a higher number of the local authorities to be performing closer to a level exhibited by London while highlighting the

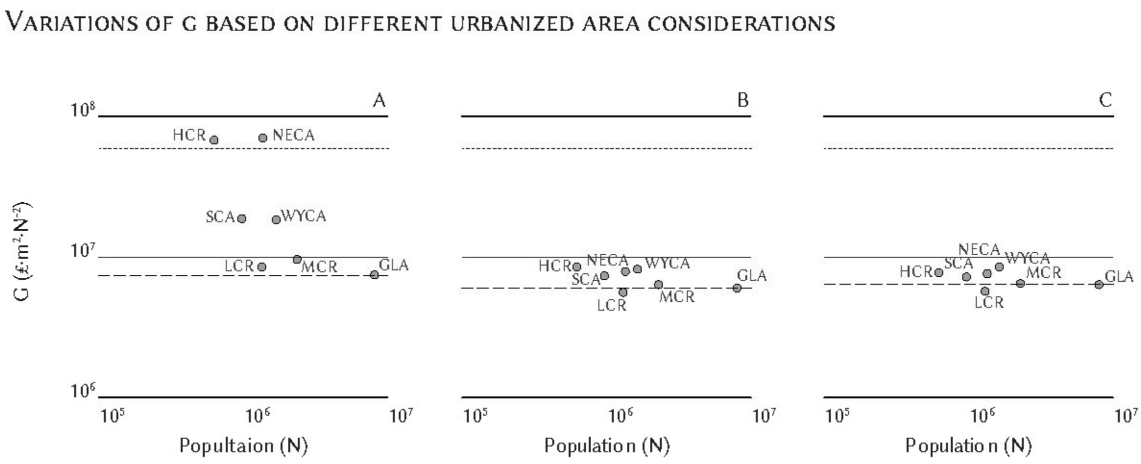


Figure B.10: Log-log plots of G against population for scenario i to iii from A to C – note that for convenience London has been assumed to act as the theoretical optimum G^* with the dotted line denoting the corresponding G_{max} limit accordingly.

comparative mobility shortcomings through smaller estimates of G in the core centers of each of the NPH city-regions.

B.2.2. Would it change our interpretations?

Throughout chapters 4 to 6, we argued that the urban network of England and Wales in general, and the Northern Powerhouse in particular, show unrealized potential, as captured by below-optimal estimates of G, due to a lack of adequate mobility regardless of the spatial scales over which cities were defined[14]. Trying to reconcile these three scenarios and our primary observations, we consider the model's size-cost formulation, $Y - W$, and the premise that settlements where $G < G^*$ are not meeting their true potential in their production of Y while those exhibiting $G > G^*$ have to pay escalating transport and mobility charges. The urban performance was estimated using the total extent of *urbanized* area within each city unit with no attention paid to overall contiguity of these patches. Under such circumstances, observations of $G < G^*$, and interpretations that suggest issues surrounding adequate mobility, are concerned with mobility of individuals over the urbanized extent having implicitly assumed no mobility tolls involved in getting people to the urbanized extent if they happen to live and interact outside these contiguous patches. Once we use total land mass area of LSOA units instead, the energy dissipated associated with the mobilization of the population rises. And so does the estimated G. Such effects are, however, minimized and less visibly evident for those units where the urbanized extent is area-filling with respect to the unit boundary chosen, e.g. local authorities of Manchester, North and South Tyneside, and the overall England and Wales.

[14] As discussed previously, an overall observation of mostly linear scaling of urban characteristics for units in England and Wales is in agreement with previous studies (Arcaute et al., 2015; Hatna, 2017).

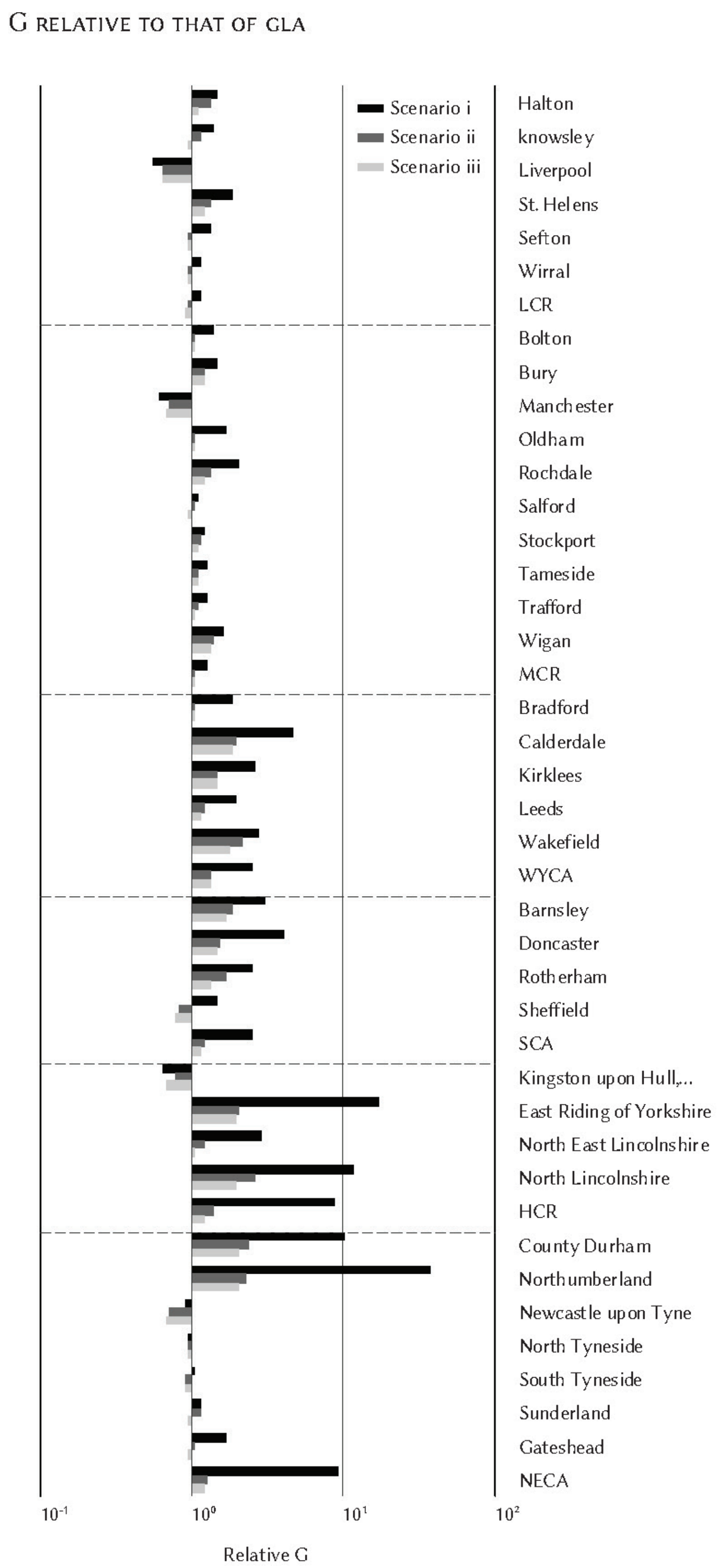


Figure B.11: Estimated relative performance ($\frac{G}{G_{GLA}}$) for the LAU1 units and the aggregated city-regions – note that the scenario i values are cut off at border for legibility.

The aggregation of local authorities into city-regions, especially under scenario ii and iii, also carries the implicit assumption that the separate distant urban parts of each member authority in a city-region are connected via a hypothetically fully efficient transportation system[15] enabling the exclusion of the uninhabited space, i.e. the non-urbanized land that in reality would increase average transportation distances, in between the core areas of local authorities as though they were connected seamlessly. This means that even in the case where we integrate the northern authorities into metropolitan regions equipped with fully capable transport infrastructure, the overall performance as a balance between the transportation costs and the economic performance would not be comparable to that of London without proper intervention to increase density, due to the relatively larger areas over which the northern authorities appear to have spread. For the units that have had an area-filling urbanized development, the findings here do, however, agree with our previous observations in suggesting that within each city-region, those local authorities which in essence house the economic hearts of the region and are more likely to be the commuter destinations for their neighbor authorities would benefit from policies that facilitate transport and mobility. This, as already suggested, is attributable to the nature of their built form relative to our choice of their bounding box, a feature they share with the England and Wales, and their lack of an extensive public transportation network similar to that the Greater London Authority enjoys.

[15] For example, electrified center-to-center passenger rail services similar to those building up the bulk of the plans for Transport for the North (2016) and Midlands Connect (2017).

In answering the question of whether or not these variations of city limits would change our previous model interpretations, it comes down to how we would/should interpret findings based on the *area* used when analyzing any given unit or a particular urban system. As we have seen, when measuring cost-size balance with the extent of urbanized area, mobility problems appears more prominent in the North of England and Wales. This is while using the total area of where we might find people living would suggest densification as required for addressing the size-cost balance because of the larger potential commuting distances that result from the region's complex urban morphology. Neither answers are demonstrably false or faulty nor are they truly contradictory when all the nuances surrounding choices of target spatial scales and units' boundaries has been explored and considered.

B.3.
Other qualifications

For the sake of completeness, it is worth us reviewing a point of technicality. We should be reminded that, although we gage the size-cost performance of cities as formulated in the social reactor model and with respect to an idealized optimal performance at G^*, we cannot truly ascertain that our urban networks or particular units do in fact maximize their $Y - W$ at our improvised and empirically estimated G^*. As briefly mentioned in chapter 3, without extended information empirically characterizing the monetary mobility costs and energetic dissipation rates, true quantification of W and hence $Y - W$ remains unfeasible. Nevertheless, the infrastructural interpretation of the $Y - W$ with respect to G remain valid when relative to a target G^* regardless of whether or not the chosen G^* satisfies $\frac{\partial(Y-W)}{\partial G} = 0$. As we saw in the previous section, one can repeat the entirety of our analysis by assuming London provides the size-cost balance to which other cities should aspire.

In concluding this appendix, we summarize our observations regarding the sensitivity of the results to potential fluctuations of city limits and estimated input data in terms of their potential effects upon conclusions we have drawn in the three pieces of work presented in chapters 4 to 6. Firstly, we showed that even relatively extreme fluctuations in the estimated economic output for individual city units does not appear to significantly change the scaling patterns previously observed in the three countries examined in this work. As such, our interpretations and findings regarding the nature of the problem, from an urban scaling perspective, affecting the urban system of England and Wales, chapter 4, or those pertaining to the macro-scale comparison of the three countries, chapter 5, continue to hold valid. We have observed that both fluctuations of the economic output and those affecting the choice of population and areal size units have the potential to change the composition of units exhibiting a need for either infrastructural intervention, i.e. deployment of better mobility and densification of the built-area. While these would not change the conclusions we have drawn in chapters 4 and 5, they will most likely alter patterns of inter-city clusterings we explored in chapter 6. The alternative examination of the northern local authorities and the stability of the estimations of G for small units with an area-filling urban extent, however, suggest that we should not expect significantly different results with respect to the importance and prominence of the short-distance pairings of units highlighting intra-city mobility as previously observed. Rather, it is the particular choice of longer-distance inter-city connections,

Figure 6.8, that would be more susceptible to such fluctuations.

Finally, the crucial point to bear in mind is that models such as the one used in this study appear to provide circumstance-specific narratives that while holding broadly valid and consistent across scales would be subject to the principle of *exact words* when matching model recommendations to a specific realization of city units. Consequently, it is important that the economic goals, say size-related productivities, are expected and monitored at the exact boundaries studied with such models and not over alternative boundaries over which similar data might be available. As we have seen, although model prediction and recommendations remain rational, changes to the boundary of the problem area also change solutions provided for that area. This brings us neatly back to the importance of a multi-scale examination of urban size-cost performance and infrastructural needs such that one can resolve results that might on the surface appear contradictory in a coherent narrative aware of the implications of boundary variations across spatial scales and alternative boundary definitions.

Bibliography

Abel, J., Dey, I., & Gabe, T. (2012). Productivity and the Density of Human Capital. *Journal of Regional Science*, *52*(4), 562–586. doi:10.1111/J.1467-9787.2011.00742.X

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Cambridge, MA: Harvard University Press.

Alonso, W. (1973). Urban Zero Population Growth. *Daedalus*, *102*(4), 191–206.

Alstott, J., Bullmore, E., & Plenz, D. (2014). Powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions. *PLoS ONE*, *9*(1), e85777. doi:10 . 1371 / JOURNAL . PONE.0085777

An, Q., An, H., Wang, L., Gao, X., & Lv, N. (2015). Analysis of Embodied Exergy Flow Between Chinese Industries Based on Network Theory. *Ecological Modelling*. Ecological Management for Human-Dominated Urban and Regional Ecosystems, *318*, 26–35. doi:10 . 1016 / J . ECOLMODEL.2015.01.020

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., & Batty, M. (2015). Constructing Cities, Deconstructing Scaling Laws. *Journal of The Royal Society Interface*, *12*(102), 20140745. doi:10.1098/RSIF.2014.0745

Arcaute, E., Molinero, C., Hatna, E., Murcio, R., Vargas-Ruiz, C., Masucci, P., & Batty, M. (2016). Cities and Regions in Britain Through Hierarchical Percolation. *Journal of The Royal Society Open Science*, *3*(4), 150691. doi:10.1098/RSOS.150691

Auerbach, F. (1913). Das Gesetz Der Bevölkerungskonzentration. *Petermanns Geographische Mitteilungen*, *59*(1), 74–76.

Barlow Commission. (1940). *Royal Commission on the Distribution of the Industrial Population*. HM Stationery Office. London, UK.

Barreira-González, P., Gómez-Delgado, M., & Aguilera-Benavente, F. (2015). From Raster to Vector Cellular Automata Models: A New Approach to Simulate Urban Growth with the Help of Graph Theory. *Computers, Environment and Urban Systems*, *54*, 119–131. doi:10 . 1016 / J . COMPENVURBSYS.2015.07.004

Batty, M. (2008). The Size, Scale, and Shape of Cities. *Science*, *319*(5864), 769–771. doi:10.1126/SCIENCE.1151419

Batty, M. (2013). *The New Science of Cities*. Cambridge, MA: MIT Press.

Batty, M., & Longley, P. (1994). *Fractal Cities: A Geometry of Form and Function*. London, UK: Academic Press.

Beaudry, C., & Schiffauerova, A. (2009). Who's Right, Marshall or Jacobs? The Localization Versus Urbanization Debate. *Research Policy*, *38*(2), 318–337. doi:10 . 1016 / J . RESPOL.2008.11.010

Bergsdal, H., Brattebø, H., & Müller, D. (2014). Dynamic Material Flow Analysis for Pcbs in the Norwegian Building Stock. *Building Research & Information*, *42*(3), 359–370. doi:10.1080/09613218.2014.887898

Bettencourt, L. M. A. (2013). The Origins of Scaling in Cities. *Science*, *340*(6139), 1438–1441. doi:10.1126/SCIENCE.1235823

Bettencourt, L. M. A., & Lobo, J. (2016). Urban Scaling in Europe. *Journal of The Royal Society Interface*, *13*(116), 20160005. doi:10.1098/RSIF.2016.0005

Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C., & West, G. (2007). Growth, Innovation, Scaling, and the Pace of Life in Cities. *Proceedings of the National Academy of Sciences*, *104*(17), 7301–7306. doi:10.1073/PNAS.0610172104

Bettencourt, L. M. A., Lobo, J., Strumsky, D., & West, G. (2010). Urban Scaling and Its Deviations: Revealing the Structure of Wealth, Innovation and Crime Across Cities. *PLoS ONE*, *5*(11), e13541. doi:10 . 1371/JOURNAL.PONE.0013541

Bettencourt, L. M. A., Lobo, J., & Youn, H. (2013). The Hypothesis of Urban Scaling: Formalization, Implications and Challenges.

Bettencourt, L. M. A., & West, G. (2010). A Unified Theory of Urban Living. *Nature*, *467*(7318), 912–913. doi:10.1038/467912A

Bettignies, Y., Meirelles, J., Fernandez, G., Meinherz, F., Hoekman, P., Bouillard, P., & Athanassiadis, A. (2019). The Scale-Dependent Behaviour of Cities: A Cross-Cities Multiscale Driver Analysis of Urban Energy Use. *Sustainability*, *11*(12), 3246. doi:10.3390/su11123246

Bibby, P., & Brindley, P. (2014). *Urban and Rural Classification of English Local Authority Districts and Similar Geographical Units in England: Methodology*. Department for Environment, Food & Rural Affairs. London, UK.

Bieber, A., Massot, M.-H., & Orfeuil, J.-P. (1994). Prospects for Daily Urban Mobility. *Transport Reviews*, *14*(4), 321–339. doi:10.1080/01441649408716888

Bircknell, A. (2018). *Sub-National Consumption Statistics*. Department for Business, Energy & Industrial Strategy, HM Government. London, UK.

Blomqvist, L., Brook, B., Ellis, E., Kareiva, P., Nordhaus, T., & Shellenberger, M. (2013). Does the Shoe Fit? Real Versus Imagined Ecological Footprints. *PLoS Biology*, *11*(11), e1001700. doi:10.1371/JOURNAL.PBIO.1001700

Blondel, V., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast Unfolding of Communities in Large Networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. doi:2008101003130400

Bly, P., Webster, F., & Pounds, S. (1980). Effects of Subsidies on Urban Public Transport. *Transportation*, *9*(4), 311–331. doi:10.1007/BF00177696

Boyden, S. (Ed.). (1981). *The Ecology of a City and Its People: The Case of Hong Kong*. Canberra, Australia: Australian National University Press.

Brenner, N., Madden, D., & Wachsmuth, D. (2011). Assemblage Urbanism and the Challenges of Critical Urban Theory. *City*, *15*(2), 225–240. doi:10.1080/13604813.2011.568717

Bristow, D., & Kennedy, C. (2012). The Energy for Growing and Maintaining Cities. *AMBIO*, *42*(1), 41–51. doi:10.1007/S13280-012-0350-X

Bristow, D., & Kennedy, C. (2013). Maximizing the Use of Energy in Cities Using an Open Systems Network Approach. *Ecological Modelling*, *250*, 155–164. doi:10 . 1016 / J . ECOLMODEL.2012.11.005

Bristow, D., & Kennedy, C. (2015). Why Do Cities Grow? Insights from Nonequilibrium Thermodynamics at the Urban and Global Scales. *Journal of Industrial Ecology*, *19*(2), 211–221. doi:10.1111/JIEC.12239

Burger, M., Meijers, E., Hoogerbrugge, M., & Tresserra, J. (2015). Borrowed Size, Agglomeration Shadows and Cultural Amenities in North-West Europe. *European Planning Studies*, *23*(6), 1090–1109. doi:10.1080/09654313.2014.905002

Burgess, E. (1925). The Growth of the City: An Introduction to a Research Project. In *The City* (pp. 47–63). Chicago, IL: University of Chicago Press.

Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart Cities in Europe. *Journal of Urban Technology*, *18*(2), 65–82. doi:10.1080/10630732.2011.601117

Centre for Cities. (2015). Northern Powerhouse Factsheet: Key Figures on the City Regions in the Northern Powerhouse.

Chandra, A., & Thompson, E. (2000). Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional Science and Urban Economics*, *30*(4), 457–490. doi:10.1016/S0166-0462(00)00040-5

Chen, G., & Qi, Z. (2007). Systems Account of Societal Exergy Utilization: China 2003. *Ecological Modelling*, *208*(2–4), 102–118. doi:10.1016/J.ECOLMODEL.2007.04.017

Chen, Y. (2010). Characterizing Growth and Form of Fractal Cities with Allometric Scaling Exponents. *Discrete Dynamics in Nature and Society*, *2010*, e194715. doi:10.1155/2010/194715

Chen, Y. (2014). An Allometric Scaling Relation Based on Logistic Growth of Cities. *Chaos, Solitons & Fractals*, *65*, 65–77. doi:10.1016/J.CHAOS.2014.04.017

Cheshire, P. (1999). Chapter 35 Trends in Sizes and Structures of Urban Areas. *Handbook of Regional and Urban Economics*. Applied Urban Economics, *3*, 1339–1373. doi:10.1016/S1574-0080(99)80004-2

Cheshire, P. (2013). Land Market Regulation: Market Versus Policy Failures. *Journal of Property Research*, *30*(3), 170–188. doi:10.1080/09599916.2013.791339

Cheshire, P., Hilber, C. A. L., & Kaplanis, I. (2015). Land Use Regulation and Productivity—Land Matters: Evidence from a UK Supermarket Chain. *Journal of Economic Geography*, *15*(1), 43–73. doi:10.1093/jeg/lbu007

Chinitz, B. (1961). Contrasts in Agglomeration: New York and Pittsburgh. *The American Economic Review*, *51*(2), 279–289.

Chopard, B., & Droz, M. (1998). *Cellular Automata Modeling of Physical Systems*. Cambridge, UK: Cambridge University Press.

Christaller, W. (1966). *Central Places in Southern Germany* (C. Baskin, Trans.). Englewood Cliffs, NJ: Prentice-Hall.

Ciccone, A., & Hall, R. (1996). Productivity and the Density of Economic Activity. *American Economic Review*, *86*(1), 54–70.

Clauset, A., Shalizi, C. R., & Newman, M. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*, *51*(4), 661–703. doi:10.1137/070710111

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection. *Econometrica*, *80*(6), 2543–2594. doi:10/f4gk8n

Coombes, M., & Office for National Statistics. (2015). *Travel-to-Work Areas* (Report No. RR2015/05). Centre for Urban & Regional Development Studies (CURDS). Newcastle, UK.

Cottineau, C. (2016). MetaZipf. (Re)Producing Knowledge About City Size Distributions.

Cottineau, C., Finance, O., Hatna, E., Arcaute, E., & Batty, M. (2018). Defining Urban Clusters to Detect Agglomeration Economies. *Environment and Plan-*

ning B, 2399808318755146. doi:10.1177/2399808318755146

Cottineau, C., Hatna, E., Arcaute, E., & Batty, M. (2015). Paradoxical Interpretations of Urban Scaling Laws.

Cottineau, C., Hatna, E., Arcaute, E., & Batty, M. (2017). Diverse Cities or the Systematic Paradox of Urban Scaling Laws. *Computers, Environment and Urban Systems*. Spatial Analysis with Census Data: Emerging Issues and Innovative Approaches, *63*, 80–94. doi:10.1016/J.COMPENVURBSYS.2016.04.006

Crescenzi, R., & Rodríguez-Pose, A. (2012). Infrastructure and regional growth in the European Union*. *Papers in Regional Science*, *91*(3), 487–513. doi:10.1111/j.1435-5957.2012.00439.x

de Groot, H., Poot, J., & Smit, M. (2007). Agglomeration Externalities, Innovation and Regional Growth: Theoretical Perspectives and Meta-Analysis. In *Handbook of Regional Growth and Development Theories*. Cheltenham, UK: Edard Elgar Publishing.

de Melo, P., Graham, D., & Noland, R. (2009). A Meta-Analysis of Estimates of Urban Agglomeration Economies. *Regional Science and Urban Economics*, *39*(3), 332–342. doi:10.1016/J.REGSCIURBECO.2008.12.002

Deilmann, C. (2009). Urban Metabolism and the Surface of the City. In *Guiding Principles for Spatial Development in Germany* (pp. 1–16). German Annual of Spatial Research and Policy. doi:10.1007/978-3-540-88839-0_7

Del Bo, C., & Florio, M. (2008). Infrastructure and Growth in the European Union: An Empirical Analysis at the Regional Level in a Spatial Framework. *University of Milan Department of Economics, Business, and Statistics departmental working papers*, *37*.

Department for Communities and Local Government. (2017). *Midlands Engine Strategy*. Department for Communities and Local Government, HM Government. London, UK.

Department for Environment, Food & Rural Affairs. (2015). 2011 Rural-Urban Classification of Local Authority Districts and Other Higher Level Geographies. *GOV.UK*.

Department for Transport. (2017). *Natinal Travel Survey: England 2016*. Department for Transport, HM Government. London, UK.

Department of Economic and Social Affairs. (2014). *World Urbanization Prospects, the 2014 Revision* (tech. rep. No. ST/ESA/SER.A/352). United Nations. New York, NY.

Diao, M., Zhu, Y., & Zhu, J. (2017). Intra-City Access to Inter-City Transport Nodes: The Implications of High-Speed-Rail Station Locations for the Urban Development of Chinese Cities. *Urban Studies*, *54*(10), 2249–2267. doi:10 . 1177 / 0042098016646686

Dorling, D. (2010). Persistent North–South Divides. In N. Coe & A. Jones (Eds.), *The Economic Geography of the UK* (pp. 12–28). London, UK: SAGE Publications Ltd.

Duranton, G., & Puga, D. (2004). Micro-Foundations of Urban Agglomeration Economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063–2117). Cities and Geography. doi:10.1016/S1574-0080(04)80005-1

Echenique, M., Hargreaves, A., Mitchell, G., & Namdeo, A. (2012). Growing Cities Sustainably. *Journal of the American Planning Association*, *78*(2), 121–137. doi:10.1080/01944363.2012.666731

Elmenreich, W., & Fehérvári, I. (2011). Evolving Self-Organizing Cellular Automata Based on Neural Network Genotypes. In C. Bettstetter & C. Gershenson (Eds.), *Self-Organizing Systems* (6557, pp. 16–25). Lecture Notes in Computer Science. doi:10.1007/978-3-642-19167-1_2

European Environment Agency. (2000). *CORINE Land Cover Technical Guide - Addendum 2000* (Technical Report No. 40). European Environment Agency. Copenhagen, Denmark.

Eurostat. (2016). GEOSTAT 2011 Grid Dataset. *Eurostat: Your Key to European Statistics.*

Eurostat. (2017). *Methodological Manual on City Statistics: 2017 Edition.* Publications Office of the European Union. Luxembourg City, Luxembourg.

Ewing, R., Haliyur, P., & Page, W. (1994). Getting Around a Traditional City, a Suburban Planned Unit Development, and Everything in Between. *Transportation Research Record*, (1466), 53–62.

Fath, B., & Borrett, S. (2006). A MATLAB® Function for Network Environ Analysis. *Environmental Modelling & Software*, *21*(3), 375–405. doi:10.1016/J.ENVSOFT.2004.11.007

Fath, B., & Patten, B. (1999). Review of the Foundations of Network Environ Analysis. *Ecosystems*, *2*(2), 167–179. doi:10.1007/S100219900067

Feng, Y., & Liu, Y. (2015). Fractal Dimension as an Indicator for Quantifying the Effects of Changing Spatial Scales on Landscape Metrics. *Ecological Indicators*, *53*, 18–27. doi:10.1016/J.ECOLIND.2015.01.020

Feng, Y., Liu, Y., & Batty, M. (2015). Modeling Urban Growth with GIS Based Cellular Automata and Least Squares SVM Rules: A Case Study in Qingpu–Songjiang Area of Shanghai, China. *Stochastic Environmental Research and Risk Assessment*, *30*(5), 1387–1400. doi:10.1007/S00477-015-1128-Z

Fisher-Gewirtzman, D., & Blumenfeld-Liberthal, E. (2012). An Agent-Based Model for Simulating Urban Morphology: Sachnin as a Case Study. *Survey Review*, *44*(325), 162–167. doi:10.1179/1752270612Y.0000000001

Florida, R., Adler, P., & Mellander, C. (2017). The City as Innovation Machine. *Regional Studies*, *51*(1), 86–96. doi:10.1080/00343404.2016.1255324

Fragkias, M., Lobo, J., Strumsky, D., & Seto, K. (2013). Does Size Matter? Scaling of CO2 Emissions and US Urban Areas. *PLoS ONE*, *8*(6), e64727. doi:10.1371/JOURNAL.PONE.0064727

Fujita, M., Krugman, P. R., & Venables, A. (1999). *The Spatial Economy: Cities, Regions, and International Trade.* MIT Press.

Fujita, M., Krugman, P., & Mori, T. (1999). On the Evolution of Hierarchical Urban Systems. *European Economic Review*, *43*(2), 209–251. doi:10.1016/S0014-2921(98)00066-X

Fujita, M., & Mori, T. (1996). The role of ports in the making of major cities: Self-agglomeration and hub-effect. *Journal of Development Economics.* Increasing Returns, Monopolistic Competition and Economic Development, *49*(1), 93–120. doi:10.1016/0304-3878(95)00054-2

Fujita, M., & Thisse, J.-F. (2002). *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth.* Cambridge, UK ; New York: Cambridge University Press.

Gabaix, X. (1999). Zipf's Law for Cities: An Explanation. *Quarterly Journal of Economics*, *114*(3), 739–767. doi:10 . 1162 / 003355399556133

Gandy, M. (2004). Rethinking Urban Metabolism: Water, Space and the Modern City. *City*, *8*(3), 363–379. doi:10 . 1080 / 1360481042000313509

Gao, L., Shan, X., Qin, Y., Yu, S., Xu, L., & Gao, Z.-Y. (2018). Scaling Tunable Network Model to Reproduce the Density-Driven Superlinear Relation. *Chaos*, *28*(3), 033122. doi:10.1063/1.5023736

Gardiner, B., Martin, R., Sunley, P., & Tyler, P. (2013). Spatially Unbalanced Growth in the British Economy. *Journal of Economic Geography*, *13*(6), 889–928. doi:10.1093/JEG/LBT003

Geary, F., & Stark, T. (2015). Regional GDP in the UK, 1861–1911: New Estimates. *The*

Economic History Review, *68*(1), 123–144. doi:10.1111/1468-0289.12061

Glaeser, E. L. (2010). Introduction to "Agglomeration Economics". In E. L. Glaeser (Ed.), *Agglomeration Economics* (pp. 1–14). National Bureau of Economic Research Conference Report. Chicago, IL: The University of Chicago Press.

Glaeser, E. L., & Gottlieb, J. (2009). The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. *Journal of Economic Literature*, *47*(4), 983–1028. doi:10.1257/JEL.47.4.983

Glaeser, E. L., Kallal, H., Scheinkman, J., & Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, *100*(6), 1126. doi:10.1086/261856

Glaeser, E. L., & Kohlhase, J. (2003). Cities, Regions and the Decline of Transport Costs. *Papers in Regional Science*, *83*(1), 197–228. doi:10.1007/S10110-003-0183-X

Glaeser, E. L., Ponzetto, G., & Zou, Y. (2016). Urban Networks: Connecting Markets, People, and Ideas. *Papers in Regional Science*, *95*(1), 17–59. doi:10.1111/PIRS.12216

Golubiewski, N. (2012a). Is There a Metabolism of an Urban Ecosystem? An Ecological Critique. *AMBIO*, *41*(7), 751–764. doi:10.1007/S13280-011-0232-7

Golubiewski, N. (2012b). The Power of Language in Feedback Metaphors: A Response to Kennedy. *AMBIO*, *41*(7), 767–768. doi:10.1007/S13280-012-0345-7

Gomez-Lievano, A., Patterson-Lomba, O., & Hausmann, R. (2017). Explaining the Prevalence, Scaling and Variance of Urban Phenomena. *Nature Human Behaviour*, *1*(1), 0012. doi:10.1038/S41562-016-0012

Gomez-Lievano, A., Youn, H., & Bettencourt, L. M. A. (2012). The Statistics of Urban Scaling and Their Connection to Zipf's Law. *PLoS ONE*, *7*(7), e40393. doi:10.1371/JOURNAL.PONE.0040393

Gonzalez de Durana, J. M., Barambones, O., Kremers, E., & Varga, L. (2014). Agent Based Modeling of Energy Networks. *Energy Conversion and Management*, *82*, 308–319. doi:10.1016/J.ENCONMAN.2014.03.018

Gordon, I. (2008). Density and the Built Environment. *Energy Policy*. Foresight Sustainable Energy Management and the Built Environment Project, *36*(12), 4652–4656. doi:10.1016/J.ENPOL.2008.09.002

Gregory, M. (2014). Sub-National Consumption Statistics: Methodology and Guidance Booklet.

Gunn, H. (1981). Travel Budgets—a Review of Evidence and Modelling Implications. *Transportation Research Part A*, *15*(1), 7–23. doi:10.1016/0191-2607(83)90012-2

Haberl, H., Steinberger, J., Plutzar, C., Erb, K.-H., Gaube, V., Gingrich, S., & Krausmann, F. (2012). Natural and Socioeconomic Determinants of the Embodied Human Appropriation of Net Primary Production and Its Relation to Other Resource Use Indicators. *Ecological Indicators*, *23*, 222–231. doi:10.1016/J.ECOLIND.2012.03.027

Hall, P., & Pain, K. (Eds.). (2012). *The Polycentric Metropolis: Learning from Mega-City Regions in Europe*. London, UK: Earthscan.

Harris, C., & Ullman, E. (1945). The Nature of Cities. *The Annals of the American Academy of Political and Social Science*, *242*, 7–17. doi:10 . 1177 / 000271624524200103

Hatna, E. (2017). Superlinear Scaling in the Urban System of England of Wales: A Comparison with US Cities.

Henderson, V. (1974). The Sizes and Types of Cities. *The American Economic Review*, 640–656.

Henderson, V. (1975). Congestion and Optimum City Size. *Journal of Urban Economics*, *2*(1), 48–62. doi:10.1016/0094-1190(75)90039-X

Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial Development in Cities. *Journal of Political Economy*, *103*(5), 1067–1090. doi:10.1086/262013

Henderson, V., & Venables, A. (2009). The Dynamics of City Formation. *Review of Economic Dynamics*, *12*(2), 233–254. doi:10.1016/J.RED.2008.06.003

Heynen, N., Kaika, M., & Swyngedouw, E. (Eds.). (2006). *In the Nature of Cities: Urban Political Ecology and the Politics of Urban Metabolism*. Questioning Cities Series. London, UK: Routledge.

HM Treasury. (2016). *Northern Powerhouse Strategy*. HM Treasury. London, UK.

HM Treasury. (2017). *Autumn Budget 2017*. HM Treasury. London, UK.

Hoekstra, A., Kroc, J., & Sloot, P. (Eds.). (2010). *Simulating Complex Systems by Cellular Automata*. Berlin, Germany: Springer.

Holl, A. (2006). A Review of the Firm-Level Role of Transport Infrastructure with Implications for Transport Project Evaluation. *Journal of Planning Literature*, *21*(1), 3–14. doi:10.1177/0885412206288905

Hoover, E. (1937). *Location Theory and the Shoe and Leather Industries*. Cambridge, MA: Harvard University Press.

Horta-Bernús, R., & Rosas-Casals, M. (2015). Obsolescence in Urban Energy Infrastructures: The Influence of Scaling Laws on Consumption Forecasting. *Journal of Urban Technology*, *22*(2), 3–17. doi:10.1080/10630732.2014.955340

House of Lords. (2017). *Connected and Autonomous Vehicles: The Future?* (Tech. rep. No. 2nd Report of Session 2016–17). London, UK.

Hoyt, H. (1964). Recent Distortions of the Classical Models of Urban Structure. *Land Economics*, *40*(2), 199–212. doi:10.2307/3144351

Hoyt, H., & United States Federal Housing Administration. (1939). *United States. Federal Housing Administration, the Structure and Growth of Residential Neighborhoods in American Cities*. Washington D.C., US: Government Printing Office.

Huang, S.-L. (1998). Urban Ecosystems, Energetic Hierarchies, and Ecological Economics of Taipei Metropolis. *Journal of Environmental Management*, *52*(1), 39–51. doi:10.1006/JEMA.1997.0157

Huang, S.-L., & Chen, C.-W. (2009). Urbanization and Socioeconomic Metabolism in Taipei. *Journal of Industrial Ecology*, *13*(1), 75–93. doi:10.1111/J.1530-9290.2008.00103.X

Huang, S.-L., & Chen, C.-S. (1990). A System Model to Analyse Environmental Carrying Capacity for Managing Urban Growth of the Taipei Metropolitan Region. *Journal of Environmental Management*, *31*(1), 47–60. doi:10.1016/S0301-4797(05)80014-3

Huang, S.-L., & Hsu, W.-L. (2003). Materials Flow Analysis and Emergy Evaluation of Taipei's Urban Construction. *Landscape and Urban Planning*, *63*(2), 61–74. doi:10.1016/S0169-2046(02)00152-4

Hui, S. (2001). Low Energy Building Design in High Density Urban Cities. *Renewable Energy*, *24*(3–4), 627–640. doi:10.1016/S0960-1481(01)00049-0

Hupkes, G. (1982). The Law of Constant Travel Time and Trip-Rates. *Futures*, *14*(1), 38–46. doi:10.1016/0016-3287(82)90070-2

Infrastructure and Projects Authority. (2015). *National Infrastructure Pipeline Factsheet*. HM Treasury. London, UK.

International Energy Agency. (2013). *World Energy Outlook 2013*. Paris, France: International Energy Agency.

Jacobs, J. (1961). *The Death and Life of Great American Cities* (Modern Library Edition). New York, NY: Modern Library.

Jacobs, J. (1970). *The Economy of Cities*. Vintage Books. New York, NY: Vintage Books.

Jones, C. (2017). Spatial Economy and the Geography of Functional Economic Areas. *Environment and Planning B*, *44*(3), 486–503. doi:10.1177/0265813516642226

Jovanovic, B., & Rob, R. (1989). The Growth and Diffusion of Knowledge. *The Review of*

Economic Studies, *56*(4), 569–582. doi:10.2307/2297501

Karathodorou, N., Graham, D., & Noland, R. (2010). Estimating the Effect of Urban Density on Fuel Demand. *Energy Economics*, *32*(1), 86–92. doi:10 . 1016 / J . ENECO.2009.05.005

Kastner, T., Erb, K.-H., & Haberl, H. (2015). Global Human Appropriation of Net Primary Production for Biomass Consumption in the European Union, 1986–2007. *Journal of Industrial Ecology*, *19*(5), 825–836. doi:10.1111/JIEC.12238

Kazanci, C., & Ma, Q. (2012). Extending Ecological Network Analysis Measures to Dynamic Ecosystem Models. *Ecological Modelling*, *242*, 180–188. doi:10 . 1016 / J . ECOLMODEL.2012.05.021

Kennedy, C., Cuddihy, J., & Engel-Yan, J. (2007). The Changing Metabolism of Cities. *Journal of Industrial Ecology*, *11*(2), 43–59. doi:10.1162/JIE.2007.1107

Kennedy, C., Pincetl, S., & Bunje, P. (2011). The Study of Urban Metabolism and Its Applications to Urban Planning and Design. *Environmental Pollution*. Selected Papers from the Conference Urban Environmental Pollution: Overcoming Obstacles to Sustainability and Quality of Life (UEP2010), 20-23 June 2010, Boston, USA, *159*(8–9), 1965–1973. doi:10.1016/J.ENVPOL.2010.10.022

Kennedy, C., Stewart, I., Ibrahim, N., Facchini, A., & Mele, R. (2014). Developing a Multi-Layered Indicator Set for Urban Metabolism Studies in Megacities. *Ecological Indicators*. Integrated Ecological Indicators for Sustainable Urban Ecosystem Evaluation and Management, *47*, 7–15. doi:10.1016/J.ECOLIND.2014.07.039

Kirby, H. (1981). Foreword. *Transportation Research Part A*, *15*(1), 1–6. doi:10.1016/0191-2607(83)90011-0

Kitamura, R., Mokhtarian, P., & Laidet, L. (1997). A Micro-Analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area. *Transportation*, *24*(2), 125–158.

Koroneos, C. J., Nanaki, E., & Xydis, G. (2011). Exergy Analysis of the Energy Use in Greece. *Energy Policy*, *39*(5), 2475–2481. doi:10.1016/J.ENPOL.2011.02.012

Krugman, P. (1991a). *Geography and Trade*. Leuven, Belgium: Leuven University Press.

Krugman, P. (1991b). Increasing Returns and Economic Geography. *Journal of Political Economy*, *99*(3), 483–499. doi:10.1086/261763

Krugman, P. (1995). Innovation and Agglomeration: Two Parables Suggested by City-Size Distributions. *Japan and the World Economy*, *7*(4), 371–390. doi:10.1016/0922-1425(95)00029-1

Kühnert, C., Helbing, D., & West, G. (2006). Scaling Laws in Urban Supply Networks. *Physica A*. Information and Material Flows in Complex NetworksInformation and Material Flows in Complex Networks, *363*(1), 96–103. doi:10.1016/J.PHYSA.2006.01.058

Laird, J. J., Nellthorp, J., & Mackie, P. J. (2005). Network effects and total economic impact in transport appraisal. *Transport Policy*, *12*(6), 537–544. doi:10.1016/j.tranpol.2005.07.003

Li, S., Zhang, Y., Yang, Z., Liu, H., & Zhang, J. (2012). Ecological Relationship Analysis of the Urban Metabolic System of Beijing, China. *Environmental Pollution*, *170*, 169–176. doi:10.1016/J.ENVPOL.2012.07.010

Lobo, J., Bettencourt, L. M. A., Strumsky, D., & West, G. (2013). Urban Scaling and the Production Function for Cities. *PLoS ONE*, *8*(3), e58407. doi:10.1371/JOURNAL.PONE.0058407

Longley, P., Batty, M., & Shepherd, J. (1991). The Size, Shape and Dimension of Urban Settlements. *Transactions of the Institute of British Geographers*, *16*(1), 75–94. doi:10.2307/622907

Lösch, A. (1954). *Economics of Location* (W. Woglom & W. Stolper, Trans.). London, UK: Oxford University Press.

Louf, R., & Barthelemy, M. (2014). Scaling: Lost in the Smog. *Environment and Planning B*, *41*(5), 767–769. doi:10.1068/B4105C

Mandelbrot, B. (2004). Fractal Geometry and Applications: Analysis, Number Theory, and Dynamical Systems. In M. Lapidus & M. van Frankenhuijsen (Eds.), *Fractal Geometry and Applications: A Jubilee of Benoit Mandelbrot* (Vol. 52). Proceedings of Symposia in Pure Mathematics. Providence, RI: American Mathematical Society.

Marshall, A. (1890). *Principles of Economics* (Reprinted 1977). London, UK: Macmillan Press.

Martin, P. (2005). The Geography of Inequalities in Europe. *Swedish Economic Policy Review*, *12*, 83–108.

Martin, R. (2008). National Growth Versus Spatial Equality? A Cautionary Note on the New 'Trade-Off' Thinking in Regional Policy Discourse. *Regional Science Policy & Practice*, *1*(1), 3–13. doi:10.1111/J.1757-7802.2008.00003.X

Martin, R., & Gardiner, B. (2017). Reviving the 'Northern Powerhouse' and Spatially Rebalancing the British Economy: The Scale of the Challenge. In C. Berry & A. Giovannini (Eds.), *Th Epolitical Economy of the Northern Powerhouse*. Basingstoke, UK: Palgrave.

Masucci, P., Arcaute, E., Hatna, E., Stanilov, K., & Batty, M. (2015). On the Problem of Boundaries and Scaling for Urban Street Networks. *Journal of The Royal Society Interface*, *12*(111), 20150763. doi:10.1098/RSIF.2015.0763

McCann, P. (2016). *The UK Regional–National Economic Problem: Geography, Globalisation and Governance*. Regions and Cities. Oxon, UK: Routledge.

McPhearson, T., Pickett, S., Grimm, N., Niemelä, J., Alberti, M., Elmqvist, T., … Qureshi, S. (2016). Advancing Urban Ecology toward a Science of Cities. *BioScience*, *66*(3), 198–212. doi:10.1093/BIOSCI/BIW002

Meijers, E. (2005). Polycentric Urban Regions and the Quest for Synergy: Is a Network of Cities More than the Sum of the Parts? *Urban Studies*, *42*(4), 765–781. doi:10.1080/00420980500060384

Midlands Connect. (2017). *Midlands Connect Strategy: Powering the Midlands Engine*. Midlands Connect. Birmingham, UK.

Miguélez, E., & Moreno, R. (2013). Skilled Labour Mobility, Networks and Knowledge Creation in Regions: A Panel Data Approach. *Annals of Regional Science*, *51*(1), 191–212. doi:10.1007/S00168-012-0526-0

Mohajeri, N., & Gudmundsson, A. (2014). Street Networks in Relation to Landforms: Implications for Fast-Growing Cities. *Journal of Geographical Sciences*, *24*(2), 363–381. doi:10.1007/S11442-014-1093-3

Mohajeri, N., Gudmundsson, A., & French, J. (2015). CO2 Emissions in Relation to Street-Network Configuration and City Size. *Transportation Research Part D*, *35*, 116–129. doi:10.1016/J.TRD.2014.11.025

Mohajeri, N., Gudmundsson, A., & Scartezzini, J.-L. (2015). Statistical-Thermodynamics Modelling of the Built Environment in Relation to Urban Ecology. *Ecological Modelling*, *307*, 32–47. doi:10 . 1016 / J . ECOLMODEL.2015.03.014

Mokhtarian, P., & Chen, C. (2004). TTB or Not TTB, That Is the Question: A Review and Analysis of the Empirical Literature on Travel Time (and Money) Budgets. *Transportation Research Part A*, *38*(9-10), 643–675. doi:10.1016/J.TRA.2003.12.004

Moomaw, R. L. (1981). Productivity and City Size: A Critique of the Evidence. *The Quarterly Journal of Economics*, *96*(4), 675–688.

Murtagh, F., & Contreras, P. (2012). Algorithms for Hierarchical Clustering: An Overview. *Wiley Interdisciplinary Reviews: Data Min-*

ing and Knowledge Discovery, *2*(1), 86–97. doi:10.1002/WIDM.53

Nash, C., Smith, A., Crozet, Y., Link, H., & Nilsson, J.-E. (2019). How to liberalise rail passenger services? Lessons from european experience. *Transport Policy*, *79*, 11–20. doi:10.1016/j.tranpol.2019.03.011

National Infrastructure Commission. (2016a). *High Speed North: A National Infrastructure Commission Report*. National Infrastructure Commission. London, UK.

National Infrastructure Commission. (2016b). *National Infrastructure Assessment: Consultation*. National Infrastructure Commission. London, UK.

NERC Environmental Information Data Centre. (2016). CORINE Land Cover 2012 for the UK, Jersey and Guernsey - Datasets. *Data.GOV.UK*.

Newcombe, K., Kalma, J., & Aston, A. (1978). The Metabolism of a City: The Case of Hong Kong. *AMBIO*, *7*(1), 3–15.

Newell, J., & Cousins, J. (2015). The Boundaries of Urban Metabolism: Towards a Political–Industrial Ecology. *Progress in Human Geography*, *39*(6), 702–728. doi:10.1177/0309132514558442

Newman, P. (1999). Sustainability and Cities: Extending the Metabolism Model. *Landscape and Urban Planning*, *44*(4), 219–226. doi:10.1016/S0169-2046(99)00009-2

Newman, P., Birrell, B., Spessa, A., Tait, D., Holmes, D., Newton, P., … Walker, B. (1996). *Human Settlements*. State of the Environment Advisory Council. Collingwood, Australia.

Newman, P., & Kenworthy, J. (1989). Gasoline Consumption and Cities. *Journal of the American Planning Association*, *55*(1), 24–37. doi:10.1080/01944368908975398

Newman, P., & Kenworthy, J. (1999). *Sustainability and Cities: Overcoming Automobile Dependence*. Washington D.C., US: Island Press.

Newton, P., Baum, S., Bhatia, S., Brown, S., Cameron, S., Foran, B., … Yenchek, D. (2001). *Human Settlements*. State of the Environment Advisory Council. Collingwood, Australia.

Niza, S., Rosado, L., & Ferrão, P. (2009). Urban Metabolism: Methodological Advances in Urban Material Flow Accounting Based on the Lisbon Case Study. *Journal of Industrial Ecology*, *13*(3), 384–405. doi:10.1111/J.1530-9290.2009.00130.X

Nordbeck, S. (1971). Urban Allometric Growth. *Geografiska Annaler. Series B, Human Geography*, *53*(1), 54–67. doi:10.2307/490887

Norman, J., MacLean, H., & Kennedy, C. (2006). Comparing High and Low Residential Density: Life-Cycle Analysis of Energy Use and Greenhouse Gas Emissions. *Journal of Urban Planning and Development*, *132*(1), 10–21. doi:10.1061/(ASCE)0733-9488(2006)132:1(10)

O'Brien, P., & Pike, A. (2015). The Governance of Local Infrastructure Funding and Financing. *Infrastructure Complexity*, *2*(1), 3. doi:10.1186/S40551-015-0007-6

O'Brien, W., Kennedy, C., Athienitis, A., & Kesik, T. (2010). The Relationship Between Net Energy Use and the Urban Density of Solar Buildings. *Environment and Planning B*, *37*(6), 1002–1021. doi:10.1068/B36030

Odum, H. (1974). Combining Energy Laws and Corollaries of the Maximum Power Principle with Visual Systems Mathematics. In *Proceedings of a Conference on Ecosystems* (pp. 239–262). Conference on Ecosystems. SIAM Institute for Mathematics and Society.

Odum, H. (1996). *Environmental Accounting: Emergy and Environmental Decision Making*. New York, NY: Wiley.

OECD (Ed.). (2012). *Redefining "Urban": A New Way to Measure Metropolitan Areas*. Paris, France: OECD.

Ohlin, B. (1933). *International and Interregional Trade*. Cambridge, MA: Harvard University Press.

Openshaw, S. (1983). *The Modifiable Areal Unit Problem*. Norfolk, UK: Geo Books.

Osborne, G. (2014). *We Need a Northern Powerhouse*.

Ottaviano, G. I. P. (2008). Infrastructure and economic geography: An overview of theory and evidence. *EIB Papers*, *13*(2), 8–35.

Overman, H., Gibbons, S., & Tucci, A. (2009). *The Case for Agglomeration Economies*. Manchester Independent Economic Review. Manchester, UK.

Palmer, J., & Cooper, I. (2011). *Great Britain's Housing Energy Fact File 2011* (tech. rep. No. 11D/866). Department of Energy & Climate Change. London, UK.

Palmer, J., & Cooper, I. (2012). *United Kingdom Housing Energy Fact File 2012* (tech. rep. No. 12D/354). Department of Energy & Climate Change. London, UK.

Parr, J. (2017). The Northern Powerhouse: A Commentary. *Regional Studies*, *51*(3), 490–500. doi:10.1080/00343404.2016.1247951

Patrício, J., Kalmykova, Y., Rosado, L., & Lisovskaja, V. (2015). Uncertainty in Material Flow Analysis Indicators at Different Spatial Levels. *Journal of Industrial Ecology*, *19*(5), 837–852. doi:10.1111/JIEC.12336

Paulsen, K. (2012). Yet Even More Evidence on the Spatial Size of Cities: Urban Spatial Expansion in the US, 1980–2000. *Regional Science and Urban Economics*, *42*(4), 561–568. doi:10.1016/J.REGSCIURBECO.2012.02.002

Persson, O. (2010). Identifying Research Themes with Weighted Direct Citation Links. *Journal of Informetrics*, *4*(3), 415–422. doi:10.1016/J.JOI.2010.03.006

Persson, O. (2014). BibExcel - A Toolbox for Bibliometricians, by Olle Persson.

Persson, O., Danell, R., & Wiborg Schneider, J. (2009). How to Use BibExcel for Various Types of Bibliometric Analysis. In F. Åstrom, R. Danell, B. Larsen, & J. Wiborg Schneider (Eds.), *Celebrating Scholarly Communication Studies: A Festschrift for Olle Persson at His 60th Birthday* (Vol. 05-S, pp. 9–24). Leuven, Belgium: International Society for Scientometrics and Informetrics.

Pidd, H. (2015). UK's North-South Divide Has Widened, Says Thinktank. *the Guardian: Cities*.

Pincetl, S. (2012). Nature, Urban Development and Sustainability – What New Elements Are Needed for a More Comprehensive Understanding? *Cities*. Current Research on Cities, *29, Supplement 2*, S32–S37. doi:10.1016/J.CITIES.2012.06.009

Porter, M. (1990). *The Competitive Advantage of Nations* (Republished with a new introduction, 1998). New York, NY: Free Press.

Pucher, J., & Buehler, R. (2008). Making Cycling Irresistible: Lessons from the Netherlands, Denmark and Germany. *Transport Reviews*, *28*(4), 495–528. doi:10.1080/01441640701806612

Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography*, *2*(4), 373–406.

Puga, D. (2010). The Magnitude and Causes of Agglomeration Economies*. *Journal of Regional Science*, *50*(1), 203–219. doi:10.1111/j.1467-9787.2009.00657.x

Puga, D., & Venables, A. J. (1997). Preferential trading arrangements and industrial location. *Journal of International Economics*, *43*(3), 347–368. doi:10.1016/S0022-1996(96)01480-8

Rae, A. (2016). The Geography of Travel to Work in England and Wales: Extracts from the 2011 Census. *Applied Spatial Analysis and Policy*, 1–17. doi:10.1007/S12061-016-9196-0

Ramaswami, A., & Chavez, A. (2013). What Metrics Best Reflect the Energy and Carbon Intensity of Cities? Insights from Theory and Modeling of 20 Us Cities. *Environmental Research Letters*, *8*(3), 035011. doi:10.1088/1748-9326/8/3/035011

Ramaswami, A., Chavez, A., & Chertow, M. (2012). Carbon Footprinting of Cities and Implications for Analysis of Urban Material and Energy Flows. *Journal of Industrial Ecology*, *16*(6), 783–785. doi:10.1111/J.1530-9290.2012.00569.X

Rauch, J. (1993). Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities. *Journal of Urban Economics*, *34*(3), 380–400. doi:10.1006/JUEC.1993.1042

Ribeiro, F., Meirelles, J., Ferreira, F., & Neto, C. (2017). A Model of Urban Scaling Laws Based on Distance Dependent Interactions. *Journal of The Royal Society Open Science*, *4*(3), 160926. doi:10.1098/RSOS.160926

Rode, P., Keim, C., Robazza, G., Viejo, P., & Schofield, J. (2014). Cities and Energy: Urban Morphology and Residential Heat-Energy Demand. *Environment and Planning B: Planning and Design*, *41*(1), 138–162. doi:10.1068/b39065

Rodríguez-Pose, A., Crescenzi, R., & Di Cataldo, M. (2018). Institutions and the Thirst for 'Prestige' Transport Infrastructure. In J. Glückler, R. Suddaby, & R. Lenz (Eds.), *Knowledge and Institutions* (pp. 227–246). Knowledge and Space. doi:10.1007/978-3-319-75328-7_11

Rosado, L., Niza, S., & Ferrão, P. (2014). A Material Flow Accounting Case Study of the Lisbon Metropolitan Area Using the Urban Metabolism Analyst Model. *Journal of Industrial Ecology*, *18*(1), 84–101. doi:10.1111/JIEC.12083

Rosenthal, S., & Strange, W. (2003). Geography, Industrial Organization, and Agglomeration. *Review of Economics and Statistics*, *85*(2), 377–393. doi:10.1162/003465303765299882

Rosenthal, S., & Strange, W. (2004). Chapter 49 Evidence on the Nature and Sources of Agglomeration Economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2119–2171). Cities and Geography. doi:10.1016/S1574-0080(04)80006-3

Rowthorn, R. (2010). Combined and Uneven Development: Reflections on the North–South Divide. *Spatial Economic Analysis*, *5*(4), 363–388. doi:10.1080/17421772.2010.516445

Rozenfeld, H., Rybski, D., Gabaix, X., & Makse, H. (2011). The Area and Population of Cities: New Insights from a Different Perspective on Cities. *American Economic Review*, *101*(5), 2205–2225. doi:10.1257/AER.101.5.2205

Rylatt, M., Gammon, R., Boait, P., Varga, L., Allen, P., Savill, M., … Strathern, M. (2013). Cascade: An Agent Based Framework for Modeling the Dynamics of Smart Electricity Systems. *Emergence: Complexity and Organization*, *15*(SPL. 2), 1–13.

Samaniego, H., & Moses, M. (2008). Cities as Organisms: Allometric Scaling of Urban Road Networks. *Journal of Transport and Land Use*, *1*(1). doi:10.5198/JTLU.V1I1.29

Schafer, A., & Victor, D. (2000). The Future Mobility of the World Population. *Transportation Research Part A*, *34*(3), 171–205. doi:10.1016/S0965-8564(98)00071-8

Schifferes, J. (2015). Economically Speaking, Britain Is Losing Three or Four Cities a Year: Blame Longer Commutes. *City Metric*.

Schläpfer, M., Bettencourt, L. M. A., Grauwin, S., Raschke, M., Claxton, R., Smoreda, Z., … Ratti, C. (2014). The Scaling of Human Interactions with City Size. *Journal of the Royal Society Interface*, *11*(98). doi:10.1098/RSIF.2013.0789

Schwarz, N. (2010). Urban Form Revisited—Selecting Indicators for Characterising European Cities. *Landscape and*

Urban Planning, *96*(1), 29–47. doi:10.1016/J.LANDURBPLAN.2010.01.007

Scott, A. (1988). *New Industrial Spaces*. Pion.

Seto, K., Güneralp, B., & Hutyra, L. (2012). Global Forecasts of Urban Expansion to 2030 and Direct Impacts on Biodiversity and Carbon Pools. *Proceedings of the National Academy of Sciences*, *109*(40), 16083–16088. doi:10.1073/PNAS.1211658109

Sim, A., Yaliraki, S., Barahona, M., & Stumpf, M. (2015). Great Cities Look Small. *Journal of The Royal Society Interface*, *12*(109), 20150315. doi:10.1098/RSIF.2015.0315

Smith, D. (2014). Domestic Energy Use in England and Wales: A 3D Density Grid Approach. *Regional Studies, Regional Science*, *1*(1), 347–349. doi:10.1080/21681376.2014.986190

Stanilov, K., & Batty, M. (2011). Exploring the Historical Determinants of Urban Growth Patterns through Cellular Automata. *Transactions in GIS*, *15*(3), 253–271. doi:10.1111/J.1467-9671.2011.01254.X

Steemers, K. (2003). Energy and the City: Density, Buildings and Transport. *Energy and Buildings*. Special Issue on Urban Research, *35*(1), 3–14. doi:10.1016/S0378-7788(02)00075-0

Stewart, H. (2015). North-South Divide Set to Widen Over Next Three Years, Study Shows. *the Guardian: Business*.

Sveikauskas, L. (1975). The Productivity of Cities. *The Quarterly Journal of Economics*, *89*(3), 393–413.

Swinney, P. (2016). *Building the Northern Powerhouse: Lessons from the Rhine-Ruhr and Randstad*. Centre for Cities. London, UK.

Swyngedouw, E. (2006). Metabolic Urbanization: The Making of Cyborg Cities. In N. Heynen, M. Kaika, & E. Swyngedouw (Eds.), *In the Nature of Cities: Urban Political Ecology and the Politics of Urban Metabolism* (pp. 20–39). Questioning Cities Series. London, UK: Routledge.

Szüle, J., Kondor, D., Dobos, L., Csabai, I., & Vattay, G. (2014). Lost in the City: Revisiting Milgram's Experiment in the Age of Social Networks. *PLoS ONE*, *9*(11), e111973. doi:10.1371/JOURNAL.PONE.0111973

Tanikawa, H., & Hashimoto, S. (2009). Urban Stock Over Time: Spatial Material Stock Analysis Using 4D-GIS. *Building Research & Information*, *37*(5-6), 483–502. doi:10.1080/09613210903169394

Terzi, F., & Kaya, S. (2008). Analyzing Urban Sprawl Patterns Through Fractal Geometry: The Case of Istanbul Metropolitan Area. *Centre for Advanced Spatial Analysis*, *WP*(144).

The Long Term Ecological Research Network. (2018). Long-Term, Broad-Scale Research to Understand Our World. *The Long Term Ecological Research Network*.

Thomas, E., Serwicka, I., & Swinney, P. (2015). *Urban Demographics: Where People Live and Work*. Centre for Cities. London, UK.

Thomlinson, R. (1969). *Urban Structure; the Social and Spatial Character of Cities* (First Edition). New York, NY: Random House.

Transport for Greater Manchester. (2017). Greater Manchester Transport Strategy 2040.

Transport for the North. (2015). *The Northern Powerhouse: One Agenda, One Economy, One North: A Report on the Northern Transport Strategy*. Department for Transport, HM Government. London, UK.

Transport for the North. (2016). *The Northern Transport Strategy: Spring 2016 Report*. Department for Transport, HM Government. London, UK.

Tsai, Y.-H. (2005). Quantifying Urban Form: Compactness Versus 'Sprawl'. *Urban Studies*, *42*(1), 141–161. doi:10.1080/0042098042000309748

United States Census Bureau. (2018). Annual Estimates of the Resident Population: April 1, 2010 to July 1, 2017. *United States Census Bureau*.

US Census Bureau Demographic Internet Staff. (2018). About Metropolitan and Micropolitan Statistical Areas. *United States Census Bureau.*

Vilhelmson, B. (1999). Daily Mobility and the Use of Time for Different Activities. the Case of Sweden. *GeoJournal*, *48*(3), 177–185. doi:10.1023/A:1007075524340

Wachsmuth, D. (2012). Three Ecologies: Urban Metabolism and the Society-Nature Opposition. *Sociological Quarterly*, *53*(4), 506–523. doi:10.1111/J.1533-8525.2012.01247.X

Wackernagel, M. (1994). *Ecological Footprint and Appropriated Carrying Capacity: A Tool for Planning Toward Sustainability* (Doctoral, The University of British Columbia, Vancouver, Canada).

Wang, J. (2016). National Spatial Strategic Plan of England. *Journal of Urban Planning and Development*, *142*(1), 04015007. doi:10.1061/(ASCE)UP.1943-5444.0000284

Warren-Rhodes, K., & Koenig, A. (2001). Escalating Trends in the Urban Metabolism of Hong Kong: 1971-1997. *AMBIO*, *30*(7), 429–438. doi:10.1579/0044-7447-30.7.429

Weaver, W. (1948). Science and Complexity. *American Scientist*, *36*, 536–544.

West, G. (2018). *Scale: The Universal Laws of Life and Death in Organisms, Cities and Companies.* London: W&N.

West, G., Brown, J., & Enquist, B. (1997). A General Model for the Origin of Allometric Scaling Laws in Biology. *Science*, *276*(5309), 122–126. doi:10.1126/SCIENCE.276.5309.122

White, H., & Griffith, B. (1981). Author Cocitation: A Literature Measure of Intellectual Structure. *Journal of the American Society for Information Science*, *32*(3), 163–171. doi:10.1002/ASI.4630320302

Wolfram, S. (1984). Cellular Automata as Models of Complexity. *Nature*, *311*(5985), 419–424. doi:10.1038/311419A0

Wolman, A. (1965). The Metabolism of Cities. *Scientific American*, *213*(3), 178–190. doi:10.1038/SCIENTIFICAMERICAN0965-178

Wolmar, C. (2016). In the Age of HS2: The Past Teaches Us There Are Far Better Ways to Run Our Railways. *the Guardian: Opinion.*

World Energy Council. (2013). *World Energy Resources: 2013 Survey* (tech. rep. No. 4184478). World Energy Council. London, UK.

Worldwatch Institute. (2007). *State of the World 2007: Our Urban Future* (L. Starke, Ed.). New York, NY: W.W. Norton.

Xie, Y., Batty, M., & Zhao, K. (2007). Simulating Emergent Urban Form Using Agent-Based Modeling: Desakota in the Suzhou-Wuxian Region in China. *Annals of the Association of American Geographers*, *97*(3), 477–495. doi:10.1111/J.1467-8306.2007.00559.X

Yakubo, K., Saijo, Y., & Korošak, D. (2014). Superlinear and Sublinear Urban Scaling in Geographical Networks Modeling Cities. *Physical Review E*, *90*(2). doi:10.1103/PHYSREVE.90.022803

Yang, Z., Zhang, Y., Li, S., Liu, H., Zheng, H., Zhang, J., … Liu, G. (2014). Characterizing Urban Metabolic Systems with an Ecological Hierarchy Method, Beijing, China. *Landscape and Urban Planning*, *121*, 19–33. doi:10.1016/J.LANDURBPLAN.2013.09.004

Yap, M., Correia, G., & van Arem, B. (2016). Preferences of Travellers for Using Automated Vehicles as Last Mile Public Transport of Multimodal Train Trips. *Transportation Research Part A*, *94*, 1–16. doi:10.1016/J.TRA.2016.09.003

Youn, H., Gastner, M., & Jeong, H. (2008). Price of Anarchy in Transportation Networks: Efficiency and Optimality Control. *Physical Review Letters*, *101*(12), 128701. doi:10.1103/PHYSREVLETT.101.128701

Zhang, J., Pourazarm, S., Cassandras, C., & Paschalidis, I. (2016). The Price of Anarchy in Transportation Networks by Estimating User Cost Functions from Actual Traffic Data. In *2016 IEEE 55th Conference on Decision and Control (CDC)* (pp. 789–794). 2016 IEEE 55th Conference on Decision and Control (CDC). doi:10.1109/CDC.2016.7798364

Zhang, Y., Xia, L., Fath, B., Yang, Z., Yin, X., Su, M., … Li, Y. (2016). Development of a Spatially Explicit Network Model of Urban Metabolism and Analysis of the Distribution of Ecological Relationships: Case Study of Beijing, China. *Journal of Cleaner Production*, *112, Part 5*, 4304–4317. doi:10.1016/J.JCLEPRO.2015.06.052

Zhang, Y., Yang, Z., & Yu, X. (2015). Urban Metabolism: A Review of Current Knowledge and Directions for Future Study. *Environmental Science & Technology*, *49*(19), 11247–11263. doi:10 . 1021 / ACS . EST . 5B03060

Index