PIN - Productivity Projects Fund

Small Project Report

Quantifying Agglomeration Productivity Potential in Long-Term Infrastructure Planning

Report prepared by:

Dr Hadi Arbabi
Jordan Pannell
Dr Stephen Hincks
Dr Giuliano Punzo

(University of Sheffield)

www.productivityinsightsnetwork.co.uk

Economic and Social Research Council

About PIN

The Productivity Insights Network was established in January 2018 and is funded by the Economic and Social Research Council. As a multi-disciplinary network of social science researchers engaged with public, private, and third sector partners, our aim is to change the tone of the productivity debate in theory and practice. It is led by the University of Sheffield, with co-investigators at Cambridge Econometrics, Cardiff University, Durham University, University of Sunderland, SQW, University of Cambridge, University of Essex, University of Glasgow, University of Leeds, and University of Stirling. The support of the funder is acknowledged. The views expressed in this report are those of the authors and do not necessarily represent those of the funders.

Contents

Introduction

Current understanding of and approaches to devising and selecting infrastructural and ultimately land-use interventions for better urban economic performance are often of limited capacity in providing long-term 'place-based' blueprints. These have been best framed by Sir David Higgins, the former Chair of HS2 Limited, as a need for an overall national transport strategy for and against which individual interventions can be constructed and appraised (Economic Affairs Committee, 2015). These issues are also echoed in the Productivity Insights Network's Infrastructure and Regional and City Productivity Debates evidence reviews. These reviews have specifically highlighted a number of gaps in our understanding and framing of the wider effects and role of infrastructure, particularly that of mobility and transport, on the productivity prospects of the UK cities. Difficulties in directly determining and measuring precise improvements in overall GVA as a result of particular transport infrastructure interventions and a lack of mechanisms/tools that would enable a more longer-term oriented planning of infrastructure targets have specifically been highlighted (Docherty & Waite, 2018; Gardiner, 2018).

The present study focuses on aspects relating to the interrelation of the effects of urban connectivity, agglomeration, and morphology on efforts seeking to increase urban and regional output through transport interventions in a UK context (National Infrastructure Commission, 2017). We attempt to assess the feasibility of quantifying a sense of productivity premium that may be associated with the spatial organization and connectivity of small-area neighborhoods and their demographic profiles within cities. Using the Sheffield council area as a testbed, this project examines ways to quantify productivity potentials, albeit not in monetary terms, in long-term planning for land-use and transport infrastructure within an agglomeration-compatible framework.

Cities as geographic networks

Infrastructural and planning efforts in addressing the UK's regional productivity divisions had until recently, particularly in the North, focused on promoting agglomeration effects through implementing inter/intra-city transport schemes (Transport for the North, 2015; Lee, 2016). The idea at the core of such transport-related interventions often relies on implicit assumptions within certain agglomeration perspectives, where size productivities are a product of a mixing population and the interactions cities host and facilitate (Glaeser & Kohlhase, 2003; Florida et al., 2017).

Emerging studies on patterns of urban scaling across systems of cities have sought to formalize and explain long-observed size-related agglomerations in and across different countries by developing explicit and implicit geographically embedded network models of cities' inhabitants and mobility infrastructure (Bettencourt, 2013; Yakubo et al., 2014; Sim et al., 2015). These works make explicit the assumption within the agglomeration-related work that average-aggregated urban economic output is a function of the number of human interactions fostered and facilitated by cities. Our focus here is, hence, on modeling the underlying spatial network within the city based on attractivity of individuals to one another as a function of their skills, distance, and the ease by which these distances can be traversed.

Long-term infrastructure planning

Within these spatially explicit frameworks, long-term blueprints of *infrastructure* can then be thought of in terms of spatial layouts of the cities' inhabitants that increase or maximize the number of interactions between individuals across the city and hence its economic output. Particularly, if given a certain labor and skills profile, the difference in the number of interactions between a city as is it exists in space and how it could potentially be (re)arranged provides a gage for the magnitude of potential productivity premium related to the city's spatial organization. In trialing the viability of such an assessment, we use the urban network model developed by Yakubo et al. (2014) within an optimization framework to maximize number of city-wide interactions as a function of inter-neighborhood distance under two constraint scenarios, see methods and data section. By altering existing neighborhood connectivity patterns, we can then derive the effective spatial organization of the city's existing neighborhood demographic and morphological characteristics (land-use and infrastructure) as an overall long-term goal against which individual interventions can be constructed.

The two scenarios presented here concern optimization of the inter-neighborhood distances such that

1. optimized distances are within the minimum and maximum of existing inter-neighborhood distances, or
2. optimized distances are within 50% of their original distance,

provided that the sum total of the length of distances remains within a tolerance limit of the sum of distances for the city's original layout. The comparison of the number of interactions before and after the optimization can then be seen as a benchmark for cities performance against their own optimal geographic arrangement and hence quantifying the productivity potential of the city solely as a function of their spatial organization.

Preliminary findings

In this section we briefly summarize the key findings and observations based on this optimization of inter-neighborhood distances in Sheffield. It is worth mentioning in advance that our findings are mostly intuitive in themselves and might appear trivial. What is of importance, however, is that these findings are independently reaffirmed by methods that can be seen as more general in their assumptions especially avoiding strong ones regarding individual behavior.

Productivity premiums and density

Denser cities perform better especially if long-range connectivity is more difficult. Existing studies already highlight the significance of population size and its density, particularly in a UK context, when considering agglomeration effects in and across urban areas (Arbabi et al., 2020). It is also easy to see that the trivial solution to optimizing neighborhood-pair distances, for maximizing individual interactions for a geographically embedded social network on a flat plane, is to stack individuals vertically to maximize density and interactions.[1] In fact, Scenario 1 specifically considers the potential productivity premium of densification, while Scenario 2 provides a measure of productivity potential locked in the spatial rearrangement of the neighborhoods. Figure 1 shows the mean road distance between neighborhoods in Sheffield and the change in distances under either scenario.

We measure the city's potential productivity premiums as the number of inhabitants' interactions, under varying influence of distance on interactions, for each scenario relative to those of the city's existing geography. Table 1 outlines the magnitude of these output premiums for the two scenarios. It can be seen that in the presence of a mobility infrastructure that assists in formation of long-range interactions, ie, interactions $\propto$ distance$^{-0.5}$, there is virtually no difference in productivity of an optimized layout and only about a 12% advantage in extreme densification. However, as long-range interaction formation becomes more difficult, ie, more

[1] Note that while such an arrangement would have the highest notional economic output due to maximized interactions, it would also have the highest congestion costs per unit area if considerations of physical infrastructure are modeled within a similar framework.

Table 1. Relative interaction/productivity of optimized scenarios to original layout.

	Scenario 1	Scenario 2
Interactions ∝ Distance$^{-0.5}$	1.12	~1.00
Interactions ∝ Distance^{-1}	1.60	1.07
Interactions ∝ Distance$^{-1.5}$	3.14	1.09
Interactions ∝ Distance^{-2}	7.85	1.15

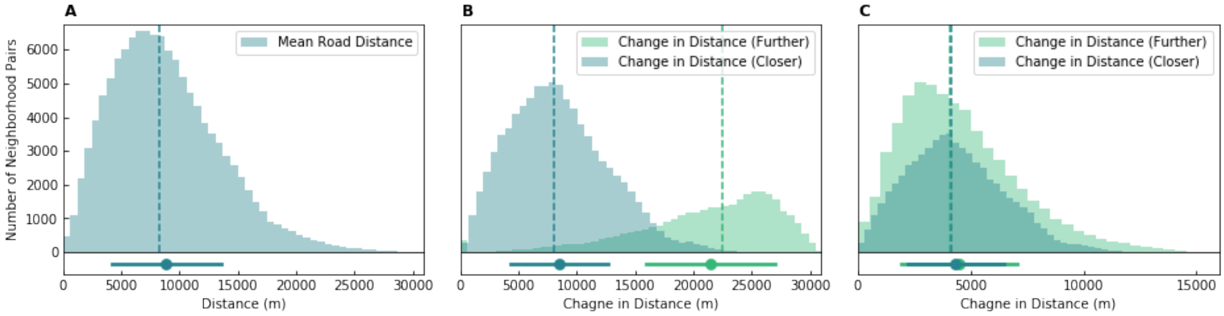

Figure 1. Histograms of neighborhood-pair mean road distances (A), change in pair distance in scenario 1 (B), and in scenario 2 (C) – dashed lines show median and pointplots mean and standard deviation of the data.

realistic provisions of mobility, simple rearrangements of the city's layout can unlock a 15% increase in output with extreme densification signaling a near sevenfold increase in output.

Distance and mixed-use planning

Homogenous deployment of mixed-used planning across the city is beneficial. For Sheffield as it exists in space, the neighborhoods with the largest number of inter-neighborhood interactions comprise the city center and exist within the city's ring road, Figure 2A. City centers often house a high-density mix of residential use, commercial activities, and crucially employment. This combined effect of their density and mixed use transforms them to the foci of both inter- and intra-neighborhood interactions. Meanwhile, the further away one gets from the city center, the more the prominence of suburban commuter-belt residential use and fewer the opportunities for local interactions.

Figures 2B and 2C show the mean of each area's inter-neighborhood distance after optimization; and as such, whether the neighborhood as whole needs to become more central to facilitate more interactions. They show that increasing the number of interactions in the city, requires a change in the land use in Sheffield such that it is more homogenously mixed-use with the current city-center type neighborhoods further away from one another and close to the rest

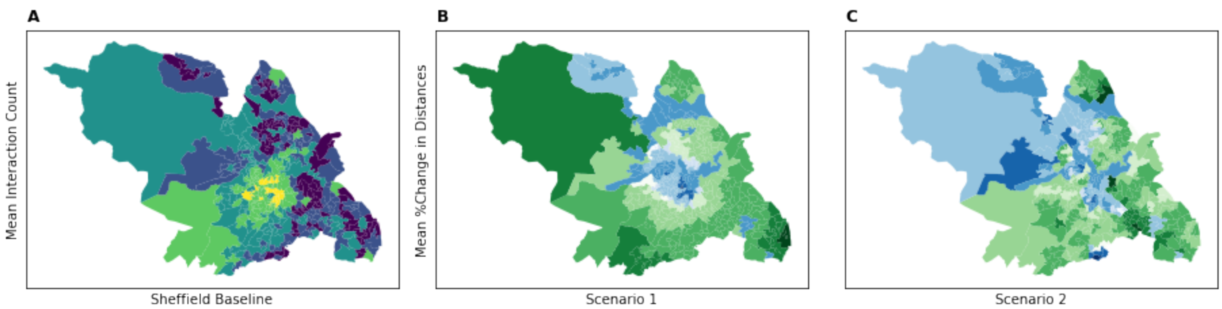

Figure 2. Choropleths showing mean population interaction count for neighborhoods in their true geography (A) and mean percentage change of pair-distances for each neighborhood under scenario 1 (B), scenario 2 (C) – in panel A, yellow and purple are highest and lowest values respectively, in panels B and C, color scheme from green to blue corresponds to an average decrease of distances to an average increase in distances.

of the neighborhoods. In essence, city-center type activities need to be more easily accessible city-wide. As such, the long-term *optimal* spatial layout of cities can be seen as one more resembling a chess board pattern of residential/commercial use that both maximizes overall city-wide interactions and facilitates walkability. This, in general, aligns with a number of planning experiments in continental Europe promoting individual clusters of neighborhoods with their own pool of employment opportunities (Speck, 2013; Hamiduddin, 2018).

For existing cities with distinct centers, this requires a breaking up of the core areas from one another and counter intuitively focusing away from the existing centers agglomeration. However, it should be noted that the density still plays a more pronounced role. The evidence, here, for more homogenously mixed spatial patterns and breaking up core areas is in addition to that of overall increased density and not a replacement for it.

Income and education

Rearrangement of neighborhoods unlocks the highest potential in relatively lower-income lower-education neighborhoods. Top 10% of the largest increases in interaction count due to the optimization of the layout involve neighborhood pairs with one area of low average income, Figure 3. More systematically, this can be seen in the correlation between mean increase in neighborhoods interactions versus their population-weighted levels of education, Figure 4. The magnitude of these effects may be due to the particularly segregated organization of neighborhoods in Sheffield (Dabinett et al., 2016), as illustrated in Figure 3A. However, the core reasoning remains the same as that underlying the need for more a homogenous mixed-use planning approach. Lower-income lower-education neighborhoods are inherently more likely to be vulnerable to effects of low long-range accessibility while simultaneously less likely to feature adequate employment opportunities. The break-up of core city-center type functionality to be

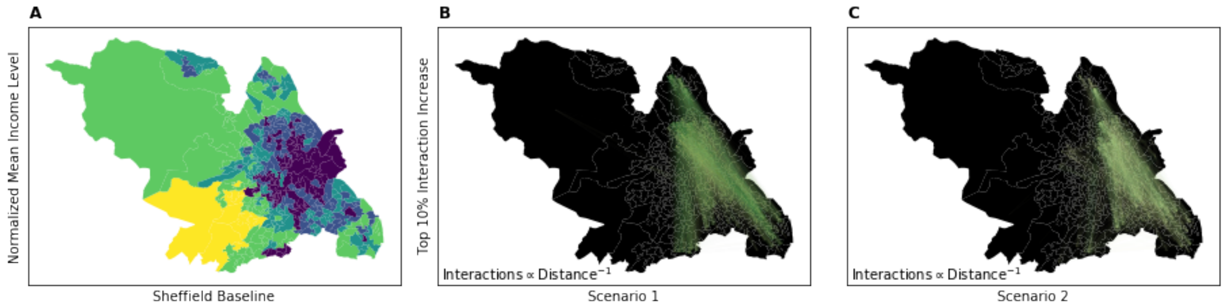

Figure 3. Choropleth showing the normalized population-weighted mean income levels (A) and the neighborhood pairs with the largest 10% increase in pair interactions for Scenario 1 (B) and Scenario 2 (C) – in panel A, yellow and purple are highest and lowest values respectively.

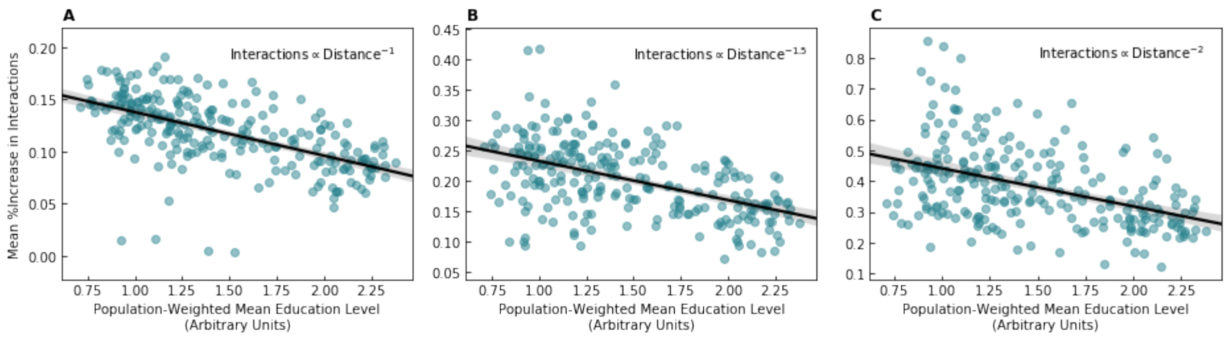

Figure 4. Mean percentage increase in neighborhood interactions against population weighted neighborhood average education levels for scenario 2 – panels A to C show results for increasing difficulty in long-range mobility/connectivity.

more evenly distributed across the city cultivates more interactions in these neighborhoods making better use of the population's otherwise spatially constrained potential.

Connectivity beyond physical mobility

Distances need not be physical. Our approach has been motivated by and particularly focused on the physical aspect of the spatial organization of cities and the provision of mobility within them. We should perhaps note that interpretations of urban scaling models of cities need not be constrained to purely physical aspects of mobility. At their core, these approaches quantify interactions and the possibility of their existence over *distances* in a city. With particular reference to the ongoing COVID-19 pandemic during the course of this study, the prevalence of home-working has forced a dramatic drop in physical intra-city mobility (Google, 2020). However, for those sectors that have remained economically active despite lockdown measures, the underlying individual interactions involved in creating cities' output has not completely vanished. Portions of these interactions are now simply forced to be made remotely. As such, spatial patterns of interaction that are identified between neighborhoods in this study, ostensibly based on road distances between them, can alternatively be interpreted as priorities for provision of alternative and/or digital connectivity infrastructure.

Methods and data

The following briefly describes the data model used in modeling the geographically embedded social network of the city.[2] Theoretical social-physical urban models work implicitly based on assumed connection between the overall number of social interactions that take place over cities and their economic output. This is often formulated as:

$$Y(N) \propto \frac{1}{2}\sum_{ij}^{N} a_{ij}y_{ij}$$

where Y(N) is the economic output of a city of population N, a_{ij} the element of the social network's adjacency matrix indicative of an interaction between individuals i and j, and y_{ij} the strength of the interaction. An interaction occurs between two individuals, ie, the adjacency matrix is 1, when a connection is deemed to have occurred subject to

$$a_{ij} = 1\, for\ \frac{x_i x_j}{l_{ij}^m} > \theta$$

where $\mathrm{l}_{\mathrm{ij}}^{\mathrm{m}}$ is given as the distance between the individuals, m representing the level of difficulty in forming long-range connections, and x_i and x_j the attractivities of two individuals. To model the attractiveness of each individual, we use neighborhood level income and education data to proxy distributions at Output Area and Lower-layer Super Output Area levels. The education data are those based on qualification levels from the 2011 census while the income data are provided by the ESRC Consumer Data Research Centre for 2016 (CDRC, 2020). For each modeled individual, the attractivity, x, is then approximated as

$$x \propto income^{-e}$$

where income is sampled for each individual from a normalized income distribution modeled at LSOA level with as a beta distribution. As a discrete variable, education levels have been

[2] The python script used for the Monte Carlo simulation and estimation of the social interactions is available from authors upon request. For more detailed description of the theoretical model used in this study, readers are directed to Yakubo et al. (2014).

Table 2. Categorized census qualification levels used determining e.

Census Qualification Level	e
No qualifications	0
Level 1 and 2 qualifications grouped together	1
Apprenticeships and level 3 qualifications grouped together	2
Level 4 qualifications and above	3

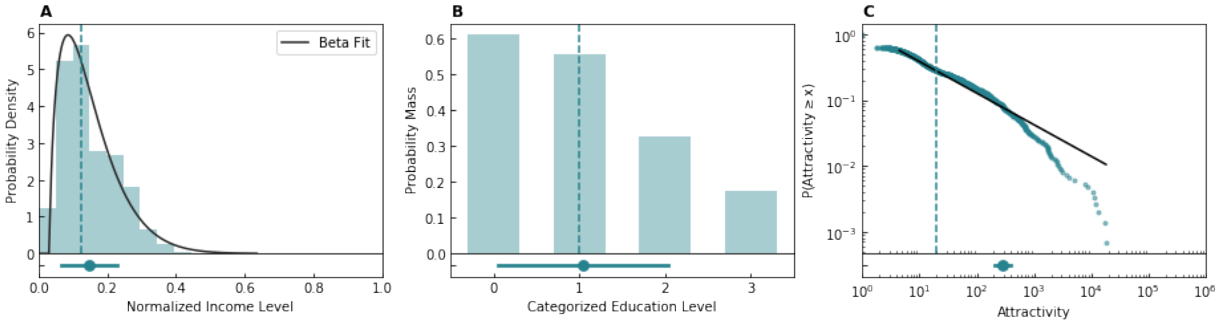

Figure 5. Plots showing distributions of normalized income levels (A), education levels (B), and complementary cumulative distribution of modeled attractivities (C) for a population of a sample neighborhood – dashed lines show median and pointplots mean and standard deviation of the data except in panel C where the pointplot shows mean and its 95%CI.

modeled by sampling from a set of population-weighted probabilities corresponding to a given qualification level at each OA, Table 2 and Figure 5.

The threshold value, θ, is given in

$$\theta = \frac{x_{min}^2}{\xi m}$$

where ξ is the threshold distance for a connection, ie, distances below which individuals connect regardless of their attractivity, whilst x^2_{min} is the minimum attractivity sampled across the overall urban population.

To calculate the total Y, we model the strength of each connection based on a power-law dependence of y_{ij} on the distance between individuals following

$$y_{ij} \propto \left(\frac{l_{ij} - \min(l)}{\max(l) - \min(l)} + 1\right)^{\eta}$$

where η is a function of the macroscale agglomeration elasticities of the larger urban system to which the city belongs,[3] and the city's fractal, D, its provision of long-range mobility, m, and the exponent of the city-wide distribution of individuals attractivity, α,[4] following

$$\eta = \begin{cases} m(\alpha - 1) - \frac{5}{6}D, & D \le m(\alpha - 1) \\ \frac{\mathrm{D}}{6}, & D > m(\alpha - 1). \end{cases}$$

In calculating mean Y, we run 2000 simulations sampling representative populations and reconstructing their connectivity. For this study, these are conducted at a LSOA level with

[3] These are observed to be roughly 7/6 in many countries including the UK (Bettencourt et al., 2007; Arbabi et al., 2019).
[4] This is highlighted in Figure 5C.

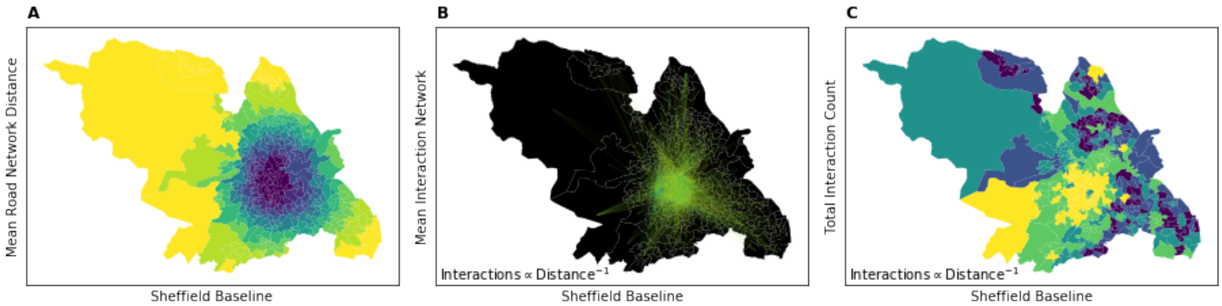

Figure 6. Choropleth showing mean road network distance (A), neighborhood pair connectivity network weighted by interaction count (B), and choropleth of neighborhood total interactions (C) for the Sheffield baseline – in panels A and C, yellow and purple are highest and lowest values respectively.

neighborhood-pair distances constructed using a closest path algorithm over the city's road network (Boeing, 2017). These consist of the estimated drive time between two randomly sampled points within corresponding LSOAs with route directionality conserved. Finally, to account for the populations at each LSOA the adjacency matrix is then multiplied by the corresponding populations of each LSOA, assuming all members in either network are fully connected. Figure 6 shows mean road network distance, sample simulated social network, and neighborhood levels of interactions for Sheffield as it exists.

The following sources of uncertainty need to be mentioned. Due to the geographical constraints of the CDRC data, the current modeling procedure assumes the attractivity distribution is homogenous per capita at a LSOA level. Given our focus on Sheffield council area, we have also assumed, unrealistically so, that no migration of labor occurs across Sheffield borders and that the labor stays within Sheffield and is not exported or imported. This, however, should not impact our preliminary finding drastically, as the estimated productivity premiums have been calculated as relative ratios. What they would influence, however, are particular patterns of optimization with respect to the spatial position of specific neighborhoods.

Acknowledgment

Figures contain National Statistics and Ordnance Survey data © Crown copyright and database right 2020.

References

Arbabi, H., Mayfield, M., & Dabinett, G. (2019). Urban performance at different boundaries in England and Wales through the settlement scaling theory. *Regional Studies*, *53*(6), 887–899. https://doi.org/10.1080/00343404.2018.1490501

Arbabi, H., Mayfield, M., & McCann, P. (2020). Productivity, infrastructure and urban density—An allometric comparison of three European city regions across scales. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *183*(1), 211–228. https://doi.org/10.1111/rssa.12490

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

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. https://doi.org/10.1073/PNAS.0610172104

Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, *65*, 126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004

CDRC. (2020). *Datasets*. ESRC Consumer Data Research Centre. https://archive.data.cdrc.ac.uk/dataset

Dabinett, G., McHugh, N., Squires, S., & Walshaw, A. (2016). *State of Sheffield*. Sheffield City Partnership. https://static1.squarespace.com/static/58d4f5f5f5e231122637f9be/t/58dbc15fdb29d6aa4b6938d0/1490796965920/State+of+Sheffield+2016.pdf

Docherty, I., & Waite, D. (2018). *Infrastructure* (PIN – 03; Evidence Review). Productivity Insights Network. https://productivityinsightsnetwork.co.uk/app/uploads/2018/07/Evidence-Review_Infrastructure-1.pdf

Economic Affairs Committee. (2015). *The Economics of High Speed 2, House of Lords Economic Affairs Committee 1st Report of Session 2014–15* (HL Paper 134). The Stationery Office Limited. https://publications.parliament.uk/pa/ld201415/ldselect/ldeconaf/134/134.pdf

Florida, R., Adler, P., & Mellander, C. (2017). The City as Innovation Machine. *Regional Studies*, *51*(1), 86–96. https://doi.org/10/f3t23g

Gardiner, B. (2018). *Regional and City Productivity Debates* (PIN – 10; Evidence Review, p. 15). Productivity Insights Network.

Glaeser, E. L., & Kohlhase, J. (2003). Cities, Regions and the Decline of Transport Costs. *Papers in Regional Science*, *83*(1), 197–228. https://doi.org/10/drgm2n

Google. (2020). *United Kingdom COVID-19 Community Mobility Report* (Community Mobility Report). Google. https://www.gstatic.com/covid19/mobility/2020-06-27_GB_Mobility_Report_en-GB.pdf

Hamiduddin, I. (2018). Journey to Work Travel Outcomes from 'City of Short Distances' Compact City Planning in Tübingen, Germany. *Planning Practice & Research*, *33*(4), 372–391. https://doi.org/10.1080/02697459.2017.1378980

Lee, N. (2016). Powerhouse of Cards? Understanding the `Northern Powerhouse'. *Regional Studies*, *0*(0), 1–12. https://doi.org/10.1080/00343404.2016.1196289

National Infrastructure Commission. (2017). *Congestion, Capacity, Carbon: Priorities For National Infrastructure* [Consultation on a National Infrastructure Assessment]. National Infrastructure Commission. https://www.nic.org.uk/wp-content/uploads/Congestion-Capacity-Carbon_-Priorities-for-national-infrastructure.pdf

Sim, A., Yaliraki, S., Barahona, M., & Stumpf, M. (2015). Great Cities Look Small. *Journal of The Royal Society Interface*, *12*(109), 20150315. https://doi.org/10.1098/RSIF.2015.0315

Speck, J. (2013). *Walkable city: How downtown can save America, one step at a time*. Macmillan.

Transport for the North. (2015). *The Northern Powerhouse: One Agenda, One Economy, One North: A Report on the Northern Transport Strategy* (p. 41). Department for Transport, HM Government. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427339/the-northern-powerhouse-tagged.pdf

Yakubo, K., Saijo, Y., & Korošak, D. (2014). Superlinear and sublinear urban scaling in geographical networks modeling cities. *Physical Review E*, *90*(2), 022803. https://doi.org/10.1103/PhysRevE.90.022803