ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2020)
183, *Part* 1, *pp.* 211–228

Productivity, infrastructure and urban density—an allometric comparison of three European city regions across scales

Hadi Arbabi, Martin Mayfield and Philip McCann

University of Sheffield, UK

[Received February 2018. Revised June 2019]

Summary. Agglomeration-based arguments citing Dutch and German city regions have been a primary driver in advocating intercity transport strategies in the north of England. We adopt an allometric urban model investigating the applicability and transferability of these transport-led agglomerative strategies promoted to address England's regional economic underperformance. This is undertaken through a comparative study of the size–cost performance balance of three city regions and the overall urban networks in the Netherlands and Germany, and England and Wales by using city units defined at different spatial scales. 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, mobility infrastructure and urban density.

Keywords: Agglomeration; Densification; Polycentric regions; Transport; Urban scaling

1. Introduction

Following an agglomeration economies line of reasoning, larger functional urban areas are thought to be associated with higher economic productivities and infrastructural efficiencies. The higher comparative productivities and efficiencies of larger cities are thought to be due to their mobility and transport cost advantages as these are instrumental by facilitating the mixing of people, ideas and goods (Glaeser, 2010). Therefore, from an agglomeration-based perspective, increased urban population and mobility are expected to enhance economic performance. These expectations have been used in support of policy arguments that champion the creation of polycentric regions through the implementation of intercity transport infrastructure. These arguments frame intercity transport as a means to increase regional economic output and productivity with the transport infrastructure of such metropolitan regions as the Dutch Randstad and the German Rhine–Ruhr given as typical examples in Europe (Burger *et al.*, 2015).

In an English context, better intercity transport links have been argued in response to the perceived productivity gap that has historically existed between the country's north and south-east. The north of England, unlike the south-east and London in particular, is comprised of cities that by international standards are suffering significant economic underperformance despite their comparable urban size (Centre for Cities, 2015). These are reported to be symptomatic of a historic regional economic performance gap that is unique to the UK (Dorling, 2010; McCann, 2016). The most recent incarnation of these infrastructural plans, the so-called 'Northern Pow-

Address for correspondence: Hadi Arbabi, Department of Civil and Structural Engineering, University of Sheffield, Sheffield, S1 3JD, UK.
E-mail: h.arbabi@sheffield.ac.uk

0964–1998/20/183211

erhouse', was launched in 2014 by the then Chancellor of the Exchequer who articulated the argument as offering the northern cities the opportunity collectively to rival global cities such as London or Tokyo by providing them with improved intercity transport links (Osborne, 2014). 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. 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 that is connected through intercity passenger rail links with decreased journey times and increased service frequency and capacity (Transport for the North, 2015; National Infrastructure Commission, 2016). Lacking from these arguments, however, has been a consideration of the compatibility of the German and Dutch case-studies when assessing such developments in the north of England.

The aim of this paper is then to explore the compatibility and transferability of such transport-driven agglomeration measures as borrowed from the Randstad and Rhine–Ruhr within an agglomeration theory compatible framework. We do so through utilizing an allometric framework adopting and applying Bettencourt's social reactor model (Bettencourt, 2013). The model enables an evaluation of the optimality of the urban size–cost performance, i.e. the balance between economic output and mobility costs that are incurred in its generation, within a system of cities. This in turn enables us to discuss the infrastructural interventions that are needed to reach this size–cost optimality and to examine the pertinence of the continental examples to the cities in the north of England. The broader contribution of the study here, however, lies in its additional focus on these questions at different geographical scales and urban boundary definitions. A further novelty of such a comparison is in its ability to facilitate an interrelated examination of economic performance and productivity, transport connectivity and mobility, and urban population and density.

As previously stated, the Randstad and Rhine–Ruhr are often cited as typical examples of productive city regions with strong intercity transport links. If the key differences underlying the higher productivities of the Randstad and Rhine–Ruhr were their intercity transport enabling such agglomeration economies, then in an allometric framework, like that of Bettencourt, we would expect distinct differences between the Northern Powerhouse and its continental comparators. Indeed, as our results will show, the overall English and Welsh urban networks do exhibit a more pronounced systemic lack of adequate mobility when compared with their Dutch and German counterparts. This could be taken to support arguments in favour of improvements to transport and mobility infrastructure as a means to boost economic productivity by enforcing increasing returns to scale for larger urban units. However, the results of our region-specific comparison highlight more nuanced differences between the three regions where the higher productivity of the Randstad and Rhine–Ruhr does not seem to be replicable in the north of England through an imitation of their intercity transport infrastructure alone.

The rest of this paper is structured as follows. The next section provides a background to urban allometry and scaling models and outlines the methods and data that were implemented in the study. This includes a summary description of Bettencourt's model derivation. (Full derivations and further discussion regarding the model can be found in Bettencourt (2013) and in Bettencourt *et al.* (2013). A brief summary derivation of Bettencourt's model, however, is provided as on-line supplementary material to this publication.) We then present a 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. Finally, a brief discussion of these national and regional comparisons and their implications are presented followed by conclusions in the last section.

The data that are analysed in the paper and the programs that were used to analyse them can be obtained from

`https://rss.onlinelibrary.wiley.com/hub/journal/1467985x/series-a-datasets`

2. Urban scaling and infrastructural needs

Recent empirical observations of population dependence of various urban characteristics are wide ranging. The consistency of these in the form of allometric power laws has prompted a notion of 'universal features' among cities (Bettencourt and West, 2010). The generic formulation of such power law relationships can be seen as

$$F(N) = F_0 N^{\beta} \tag{1}$$

or alternatively log-transformed as

$$\ln\{F(N)\} = \ln(F_0) + \beta \ln(N) \tag{2}$$

where F denotes any urban indicator of choice, e.g. economic output, urbanized area and carbon dioxide emissions, F_0 a prefactor describing the baseline prevalence of the indicator, N the urban population count and β the scaling exponent determining the scaling regime. Empirical evidence from the American, Chinese and German urban networks points to recurring values of β whereby infrastructural indicators, e.g. urbanized area and length of roads, grow sublinearly with population, $\beta \approx \frac{5}{6}$, whereas those representing wealth and information, e.g. gross domestic product, exhibit superlinear regimes, $\beta \approx \frac{7}{6}$ (Bettencourt *et al.*, 2007).

Taking shape on the basis of these empirical observations and gaining wider traction is a new science of cities that has sought to codify these observations in the form of structural and/or statistical models (Batty, 2012). There are several allometric urban models explaining empirical observations and deriving power law relationships. These range from those using probabilistic considerations of urban population and their characteristics (Gomez-Lievano *et al.*, 2016) to network conceptualizations of urban population and their connectivity embedded geographically (Yakubo *et al.*, 2014; Sim *et al.*, 2015). As previously mentioned, among the various existing urban scaling models, Bettencourt's social reactor model includes an explicit consideration of mobility effects and size–cost balances. We first outline the model's set-up before providing a concise description of the input data that were used.

2.1. Bettencourt's social reactor model

In setting up an idealized scaling model of cities, Bettencourt (2013) started from four simple assumptions:

(a) the average aggregate socio-economic product is a linear function of the sum of all local interactions (Jones, 2016);
(b) urban population is mixing uniformly and each individual has the minimum resources that are needed to travel and experience the city fully (Glaeser and Kohlhase, 2003);
(c) individual baseline production is bounded and is not a function of city size (Szüle *et al.*, 2014);
(d) 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).

(Although the first two assumptions may appear contentious, the first is supported by current empirical observations and generally agreed on across other urban scaling models (Yakubo *et al.*, 2014; Sim *et al.*, 2015; Gomez-Lievano *et al.*, 2016) whereas 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.)

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

$$Y = \bar{g} a_0 l \frac{N^2}{A_n} \tag{3}$$

where Y is the average economic output, N^2/A_n the density to the upper limit of total encounters possible ($N(N-1) \approx N^2$ for large populations) over the urbanized area A_n, a_0 and l the average effective interaction cross-section and travel path of an individual respectively, and hence the average effective area, and $\bar{g}$ the average encounter output. 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 ($\mathrm{d}G/\mathrm{d}N \approx 0$). The second assumption then derives a generic scaling for cities' volumetric area by equating *per capita* mobility costs, i.e. cost of travel, and *per capita* economic output, i.e. minimum resources for travel. Bettencourt additionally developed a scaling relationship for the energy that is dissipated over the urbanized area A_n, moving the population, goods and services, and enabling the generation of Y by treating the infrastructure network as parallel resistors. Put together, the four assumptions result in

$$\left.\begin{aligned} Y(N) &= Y_0 N^{1+H/\{D(D+H)\}}, \\ W(N) &= W_0 N^{1+H/\{D(D+H)\}}, \\ A_n(N) &= A_{n0} N^{1-H/\{D(D+H)\}} \end{aligned}\right\} \tag{4}$$

where Y, A_n and W are the average expected economic output, urbanized area and mobility costs respectively, Y_0, A_{n0} and W_0 the baseline prevalence of Y, A_n and W all functions of G, and N the population size. As can be seen, the exponents $1 \pm H/\{D(D+H)\}$ are functions of the city geometry D and the geometry of the average individual's path H. This in effect means that the exponents characterize, by proxy, the average level of mobility and accessibility across the urban network. (Note that the formulations in equation (4) represent the average expected values describing the urban behaviour across an entire urban network. For the formulation to be exact the inclusion of a fluctuation term is required (Bettencourt and Lobo, 2016). Most empirical studies, however, observe the statistics of such fluctuations to be Gaussian and zero centred for the log-transformed equation (2) for a range of urban indicators (Bettencourt *et al.*, 2007; Gomez-Lievano *et al.*, 2012).)

Imposing real geometric constraints puts the fractal dimension of the city, D, somewhere in the range [2,3]. Similar considerations would result in the geometric dimension of the travel path, H, to be confined to $[0, D)$, resulting in a range of $[0, \frac{1}{4})$ for $H/\{D(D+H)\}$. As such, in agreement with the agglomeration theory, the model expects increasing output productivities and infrastructural efficiencies for larger cities, i.e. a superlinear scaling of Y and sublinear scaling of A_n. In developing a theoretical and idealized approximation of urban networks, city geometry can be taken to be two dimensional, $D=2$, whereas 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 superlinear scaling for economic output with the exponent $\beta_Y=\frac{7}{6}$ and a sublinear 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 USA, East Asia and Europe (Bettencourt *et al.*, 2007; Bettencourt, 2013; Bettencourt and Lobo, 2016). Furthermore, since these elasticities are increasing functions of H, a lack of adequate mobility and access diminishes superlinear and sublinear effects, resulting in close-to-linear exponents. Such inadequate levels of mobility, $H<1$, can be seen as mobility patterns where individuals' access is limited and constrained to disconnected patches within the city. Finally, Bettencourt formalized the urban size–cost performance balance as the economic output less its mobility costs, $Y-W$. 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; Fig. 1. (Beware that the schematic curve that is 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. See the on-line supplementary material for an expanded expression of $Y-W$ in terms of G.)

As can be seen, the $(Y-W)$-balance grows for increasing values of G in the range $[0, G^*)$, reaching its maximum at G^*. However, for increasing values of human production beyond G^* the cost–size balance shrinks resulting in an increasingly unstable city as the costs that are associated with the mobility processes overwhelm the economic success of the city such that for $G>G_{\max}$ the city would break down to smaller functional urban zones. Bettencourt posited that, given an urban network with a relatively large number of cities, we would expect to find the statistics of G estimated for all cities to hover close to G^* as cities strive to maintain an optimal cost–size balance. (Empirical demonstrations of this for the American urban network can be found in Bettencourt (2013).) Additionally, referring 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. improvement to mobility and accessibility, enabling more urban interaction and hence higher economic output. In contrast, for cities where the economic success of the city has resulted in larger-than-desired urbanized expansion, $G>G^*$, 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.

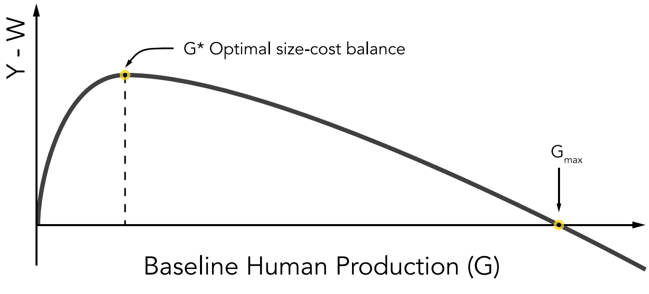


Fig. 1. Schematic illustration of cost–size balance $Y-W$, as a function of the baseline human production G

2.2. Scaling normalization and intersystem comparison

As outlined in the previous section, the overall status of mobility in an urban system and infrastructural needs of individual cities can be gleaned by investigating the population scaling of economic output and urbanized area across cities within the same connected urban system. An empirical estimate of the baseline human production for each city can easily be estimated through a rearrangement of equation (3):

$$G_i = \frac{Y_i A_{ni}}{N_i^2} \tag{5}$$

where G_i is the human production estimated for city i and Y_i, A_{ni} and N_i are the economic output, urbanized area and population of city i respectively. 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 estimate these fully, a systemwide average G^* can be obtained by substituting the scaling expressions of Y and A_n in equation (5):

$$G^* = \frac{Y_0 N_i^{1+H/\{D(D+H)\}} A_{n0} {A_{ni}}^{1-H/\{D(D+H)\}}}{N_i^2} = Y_0 A_{n0} \tag{6}$$

where Y_0 and A_{n0} are the systemwide prevalence of economic output and urbanized area respectively. Estimating an idealized optimal G^*, however, 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 (4) by using Bettencourt's theoretical ideal scaling exponents of $\beta_Y = \frac{7}{6}$ and $\beta_{A_n} = \frac{5}{6}$.

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 the indicators by using the idealized prevalence of the indicators in each system with this y-translation taking the form

$$\begin{aligned} \ln(Y_i^{\mathrm{T}}) &= \ln(Y_i) - \ln(Y_0^*) = \ln(Y_0^*) + \beta_Y \ln(N_i) + \xi_{Y\,i} - \ln(Y_0^*) = \beta_Y \ln(N_i) + \xi_{Y\,i}, \\ \ln(A_{ni}^{\mathrm{T}}) &= \ln(A_{ni}) - \ln(A_{n0}^*) = \ln(A_{n0}^*) + \beta_{A_n} \ln(N_i) + \xi_{A_n\,i} - \ln(A_{n0}^*) = \beta_{A_n} \ln(N_i) + \xi_{A_n\,i} \end{aligned} \tag{7}$$

where Y_i^{T} and A_{ni}^{T} are the normalized economic output and urbanized area for city i respectively, Y_0^* and A_{n0}^* the idealized fixed gradient systemwide prevalence of output and urbanized area respectively, and $\xi_{Y\,i}$ and $\xi_{A_n\,i}$ the fluctuation terms from the theoretical scaling for city i. Through this translation, the theoretical model of economic 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 1. The normalization enables both a comparison of size–cost performance and a multisystem examination of the population scaling by investigating power law fits to the combined data sample of the various urban networks.

2.3. Urban boundary definition

To study the urban performance balance and infrastructural needs in Germany, DE, the Netherlands, NL, and England and Wales, EW, we first need to obtain estimates for population, output and urbanized area indicators. (Files including population, output and urbanized area estimations for each boundary are available from `https://rss.onlinelibrary.com/hub/journal/1467985x/series-a-datasets` along with Jupyter notebooks containing some further insights and comments.) To estimate population at different scales we use the GEOSTAT population grid (Eurostat, 2016a), which provides population counts for the year 2011 over square grids of 1 km × 1 km area, as building blocks. CORINE land cover data are also used to estimate urbanized area for the same time interval (Copernicus Land Monitoring Service, 2016). Regional gross value-added (GVA) data for the year 2011, which are available through Eurostat (Eurostat, 2016b), are also used for the economic output indicator. The GVA data are, however, only available aggregated for the nomenclature for units of territorial statistics (NUTS) level 3 administrative boundaries. We use the Organisation for Economic Co-operation and Development's (OECD's) simplified geographical information system based method (Organisation for Economic Co-operation and Development, 2012) to break down the GVA values at NUTS level 3 to the GEOSTAT population grid based on an area- and population-weighted approach according to

$$Y_{\text{cell}} = \sum_i \frac{Y_{\text{NUTS3}} (N_{\text{cell}} / A_{\text{cell}}) A_i}{N_{\text{NUTS3}}} \tag{8}$$

where Y_{cell} denotes the total GVA share assigned to a grid cell, N_{cell} and A_{cell} the total population and area of the cell (often approximately 1 km^2 unless belonging to a coastal or border grid cell) respectively, A_i the area of the ith segment intersected by a given NUTS level 3 unit with GVA and population of Y_{NUTS3} and N_{NUTS3} respectively. The grid level GVA values are then summed back up to estimate aggregate values for the other boundaries by using a reversal of equation (8). We should acknowledge that estimating population and GVA through these area-based proportionalities is simplistic and assumes a uniform population density distribution. This could potentially result in erroneous estimates when aggregating back up to urban units that are not significantly larger than the initial grid cells (Smith, 2014). However, in the absence of data sets of better quality and/or resolution, the approach remains one of very few available options. See the on-line supplementary material for a discussion of potential implications of variations in the GVA estimates.

We adopt a mix of density-based, administrative and functional boundary definitions; Table 1. The density-based boundaries have been assembled by using the city clustering algorithm as per Rozenfeld *et al.* (2011) by merging neighbouring GEOSTAT cells with population densities above a set cut-off, e.g. 100 people km^{-2} for C100. The two boundaries representing the functional urban areas, i.e. urban audit and OECD's functional urban areas, both delineate urban areas based on considerations of the percentage of population living and working within the same area by using commuter data (Organisation for Economic Co-operation and Development, 2012; Eurostat, 2017) with the difference in the cut-off values used for population ratios and the minimum population of units. The functional urban boundaries are subject to a minimum population cut-off by definition limiting the units considered to those which are the most populated and hence urban. The raw density-based units that are created through the city clustering algorithm, however, could potentially include a large number of sparsely populated units. Instead of applying an arbitrary minimum population cut-off for these density-based boundaries, we employ the method that was described in Clauset *et al.* (2009) to estimate a lower bound for population in each density-based boundary N_{min}-values indicated in Table 1. (The Python package that was used is available in Alstott *et al.* (2014) and complementary cumulative distribution functions highlighting the population cut-offs and the approximate power law distributions for the density-based boundaries can be found in the Jupyter notebooks.) These lower bound values correspond to the values above which a coherent power law distribution, *à la* those empirically observed by Auerbach and attributed to Zipf (Auerbach, 1913; Gabaix, 1999), can be assumed to apply to the population distribution across the urban system.

Table 1. Summary of the urban boundary definitions†

Boundary		*Number of units*			N_{min}			*Number of units* $(N > N_{\text{min}})$		
		DE	*NL*	*EW*	*DE*	*NL*	*EW*	*DE*	*NL*	*EW*
C100‡	Density based	10358	634	2867	9769	4455	3895	700	235	587
C350		10072	961	2928	7847	7119	7627	965	246	481
C500		8325	957	2475	8405	6801	59698	879	255	104
C750		6117	884	2021	9317	6192	57698	768	272	112
C1000		4729	779	1692	8209	5582	55031	827	296	120
C1400		3370	649	1435	8801	4334	67495	717	339	97
NUTS level 3§	Administrative	402	40	125	34119	49364	69909	402	40	125
URBAUD§§	Functional areas	94	34	83	57161	59589	77170	94	34	83
OECD*		24	5	13	527268	692953	536892	24	5	13

†The noticeable differences in the magnitude of the population cut-offs estimated for the three countries when considering a number of the density-based boundary definitions are reflective of the population domains over which a single power law rank size distribution is coherent. DE and NL systems appear to follow such distributions over a larger portion of their smaller-sized units in contrast with the EW system where a clear shift in the distribution exponent takes place over larger population sizes; see the on-line supplementary material for distribution figures. Values of N_{min} indicated for NUTS level 3, urban audit (URBAUD) and OECD boundaries represent the population of the smallest unit rather than a cut-off used by the authors; see the on-line supplementary material.

‡Numbers indicate the minimum density value used as the cut-off when applying the city clustering algorithm to the population grid.

§Note that, for units in England and Wales, constituting members of the Greater London Authority have been merged and used as one single area.

§§Urban audit functional urban areas 2011–2014.

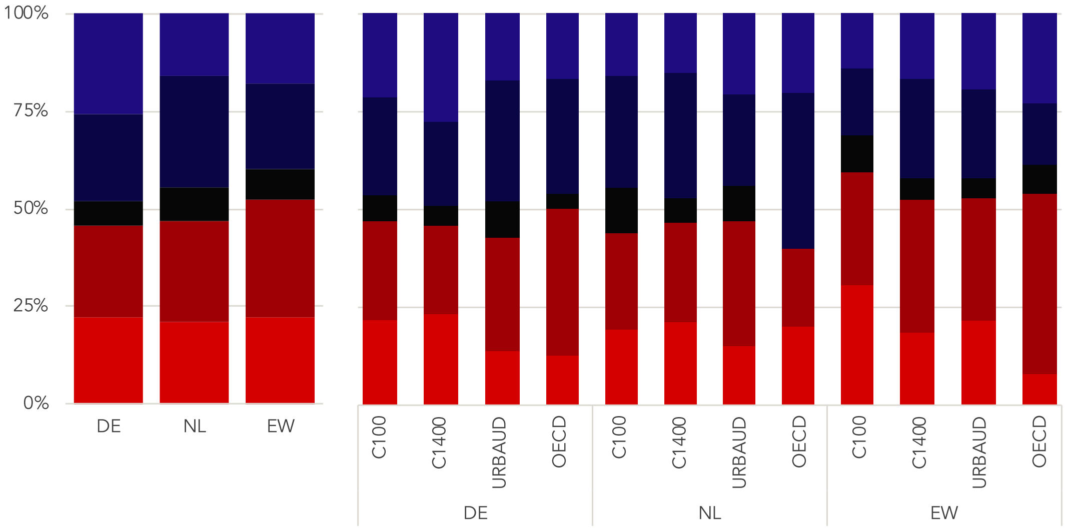
*OECD functional urban areas.

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

We begin by examining the existence of power law scaling and the empirical proximity of each country's urban network with Bettencourt's theoretical ideal. Fig. 2 shows the OLS estimates for the GVA and urbanized area scaling exponents for each boundary and country. (Use of simple OLS estimators is justified following the prior assumption and empirical observations that the scaling deviation term ξ follows a normal distribution centred on zero.) As can be seen, the scaling of urbanized area and economic output do overall display a coherent sublinear and superlinear relationship with population respectively, regardless of the choice of country and/or urban network boundary definition. The extent of sublinearity and/or superlinearity of the relationships, i.e. the strength of agglomeration effects in economic output and urbanized area, however, does vary across countries and boundary definitions. In this context, Germany shows on average the largest systemwide agglomeration elasticities for economic output followed by the Netherlands and then England and Wales. From the perspective of Bettencourt's model, the deviations from the ideal exponents of $\beta_Y = \frac{7}{6}$ and $\beta_{A_n} = \frac{5}{6}$ towards 1 indicate, on average, a systemwide lack of mobility, $H < 1$, 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 theoretical ideal for the urban audit (URBAUD) and OECD functional urban areas which are the most directly compatible boundaries to those assumed within the model's assumptions (a) and (b) (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 $\mathrm{d}G/\mathrm{d}N \approx 0$ (with the R^2 of G against N averaging around 0.03 across different

(a) (b)

Fig. 2. Plots showing the OLS estimated scaling exponents for each boundary (tabulated OLS estimates and confidence intervals for all boundaries are available in the on-line supplementary material) (– – –, theoretically ideal values for $D = 2$ and $H = 1$; ■, DE; ◆, NL; ▲, EW; ——, combined): (a) GVA exponent β_Y; (b) area exponent β_{An}



(a) (b)

Fig. 3. Bar charts displaying (a) the percentage of city units in each country and (b) for each boundary definition in the indicated range of η (bar charts for the remaining boundaries can be found in the on-line supplementary material): ■, $\eta < -0.2$; ■, $-0.2 \leqslant \eta < -0.02$; ■, $-0.02 \leqslant \eta \leqslant 0.02$; ■, $0.02 < \eta \leqslant 0.2$; ■, $0.2 < \eta$

boundaries and countries), suggests that the model's third assumption also holds; see the on-line supplementary material.

Similarly, 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; Fig. 3. Fig 3 shows the percentage of units within a given comparative performance band, $\eta\,(\equiv \ln(G/G^*))$, where increasingly negative values indicate an increasing need for better intraunit mobility and transport and larger positive values an increasing need for built area densification. It can be gleaned from the bar charts that the size–cost performance appears more symmetrically distributed around the idealized optimum, $-0.02 \leqslant \eta \leqslant 0.02$, when considering the aggregated distribution of performance balance for Germany and the Netherlands compared with those of England and Wales. (We use an arbitrary range rather than the absolute $\eta = 0$ when interpreting optimality allowing for minor variations about the empirically designated G^*.) 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 intraurban transport solutions regardless of spatial scales.

A combined interpretation of the comparative size–cost performance distribution and the overall scaling exponents that were 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. Meanwhile, the England and Wales region is further burdened with an additional prevalence of inadequate intraurban access and mixing that appears unique in its spatial persistence despite its similar exponent estimates to those of the Netherlands. For completeness, it is worth clarifying that this comparison is a comparison of the comparative agglomerative productivities gauging the increased benefits that are associated with increased size. The comparison 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 cumulative number and population of cities in each country.

4. Rhine–Ruhr, the Randstad and the 'Northern Powerhouse'

The current infrastructure plans in England and Wales, as previously mentioned in the background, focus heavily on the implementation of an intercity passenger rail solution. Combined with improving journey times and frequency, these measures have explicitly been borrowed from the Dutch Randstad and German Rhine–Ruhr to connect and transform a handful of the country's northern cities into a virtual city of a larger effective size (Infrastructure and Projects Authority, 2015; Transport for the North, 2016). The results that were presented in Section 3, 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 EW's isolated scaling and as a comparison relative to the performance of the German and Dutch urban networks. This national comparison, however, would not necessarily justify the appropriation of an explicitly intercity mobility solution from Rhine–Ruhr and the Randstad for implementation in the Northern Powerhouse. It is also crucial to note here that this examination of η masks individual economic productivity and infrastructure efficiency performance. Since η considers only the overall balance of $Y - W$, it is possible for cities to compensate for deviations from ideal scaling in one indicator, say Y, through similar deviations in the other, i.e. A_n. In such a way, considering equation (7), a city unit with lower than ideally expected economic output, $\xi_{Y_i} < 0$ for $\beta_Y = \frac{7}{6}$, can compensate by incorporating a larger effective urbanized area, $\xi_{A_{n_i}} > 0$ for $\beta_{A_n} = \frac{5}{6}$, to keep the overall G close to optimality. This leads to cities where despite a balanced size–cost performance economic underperformance may still be prevalent when compared with others.

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 deviation from the idealized expectations of economic output and urbanized area. We also examine the overall city regions that these units belong to by considering the hypothetical cities of their combined size summing their population, economic 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), corresponds to an aggregation of NUTS level 3 administrative units and is also representative of the *planned* Northern Powerhouse in England and Wales. The other is based on the extent demarcated by the largest contiguous C100 units in each region; Fig. 4. It is interesting that there is good agreement in the geography of the Randstad and Rhine–Ruhr defined either administratively or through urban proximity, i.e. the single largest contiguous unit at a threshold of 100 people km^{-2}. This, however, is not so for the realizations of the Northern Powerhouse.

4.1. A regional comparison

Proceeding with our results, Fig. 5 compares the size–cost performance of each region aggregated from units at each boundary definition and its overall deviation from the idealized economic output and urbanized area scaling. The broken diagonal line 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 Fig. 3. Comparing only the size–cost performance of the regions, not much difference could be discerned between the Randstad, Rhine–Ruhr and the Northern Powerhouse. The majority of their different realizations indicate a need for better internal mobility and mixing regardless of the choice of boundary definition or their overall extent. This is in spite of the existing intercity passenger rail infrastructure in the Randstad and Rhine–Ruhr. Out of the three, however, the Randstad shows a larger qualitative variation in estimates depending on the choice of boundary definition with the realization comprised of

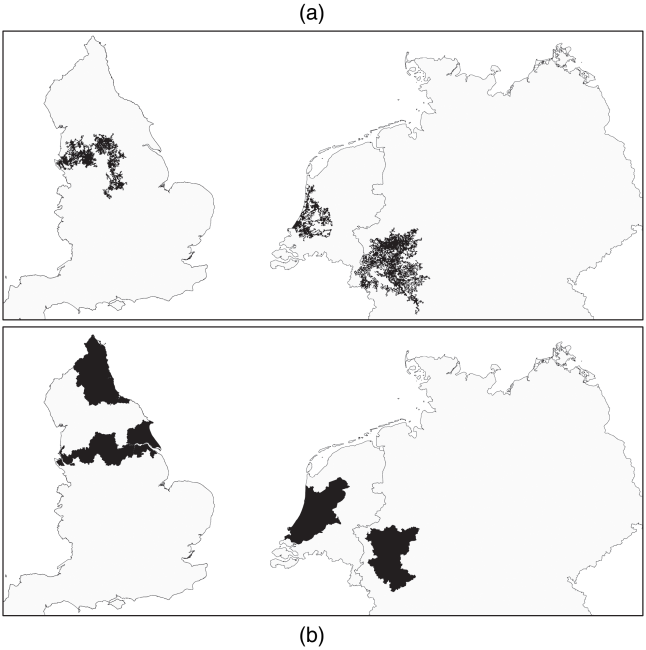
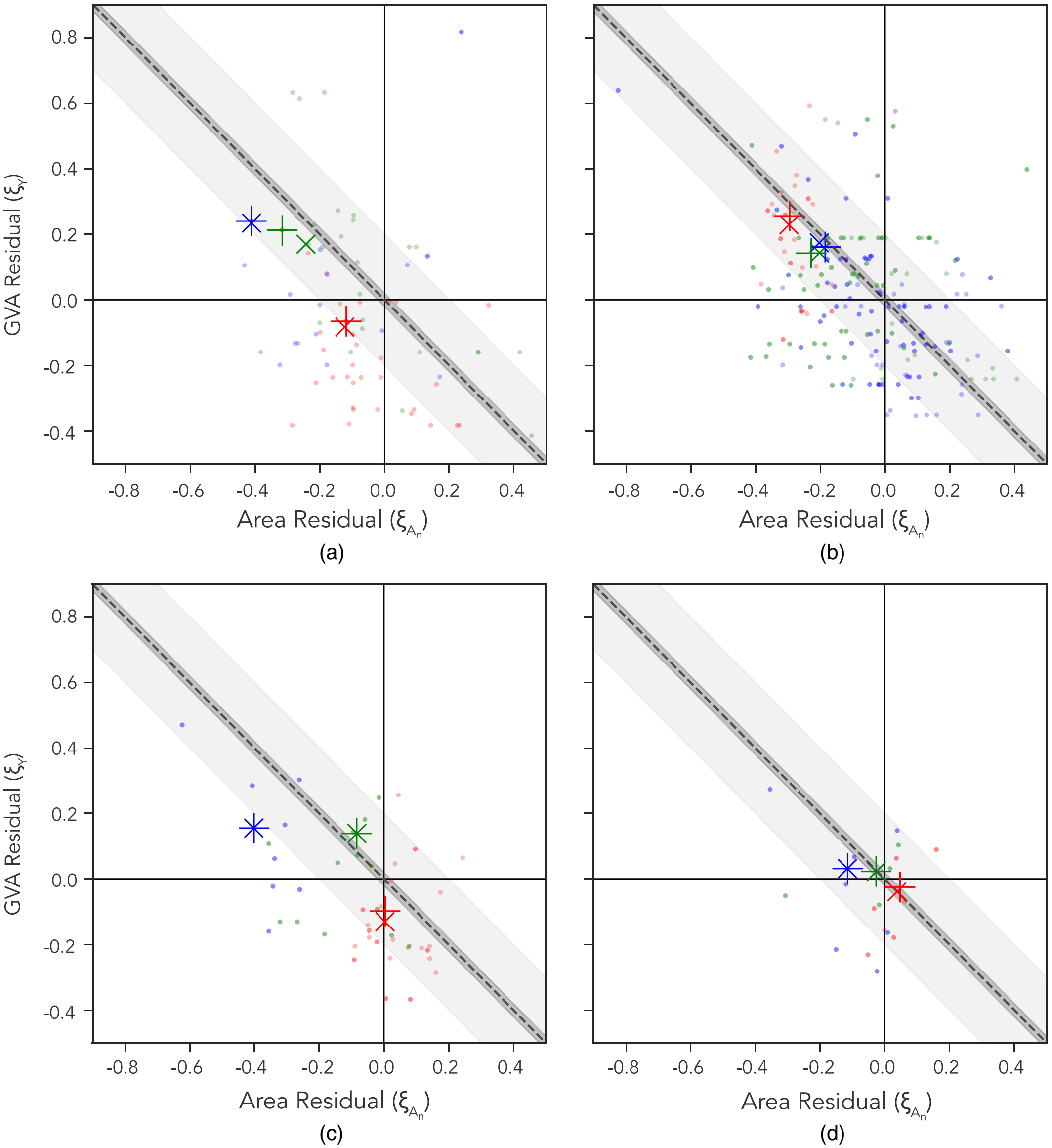


Fig. 4. Maps showing the areal extents used for allocating units to the city regions: (a) contiguous C100 units and (b) NUTS level 3 units



(a) (b) (c) (d)

Fig. 5. Scatter plot of output residual against urbanized area residual for the Randstad, Rhine–Ruhr and Northern Powerhouse assembled from units at different boundary definitions (R^2-values are calculated for the smaller markers which are indicative of the spread of the units building up the three metro regions; see the on-line supplementary material for a coloured version) (×, based on NUTS level 3 units; +, based on C100 extent; ●, DE; ●, NL; ●, EW): (a) C100, $R^2 \sim 0.20$; (b) C1400, $R^2 \sim 0.12$; (c) URBAUD, $R^2 \sim 0.26$; (d) OECD, $R^2 \sim 0.02$

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. In contrast, no rendition of Rhine–Ruhr exhibits $\eta \geqslant 0$. Meanwhile, the comparison suggests 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 that

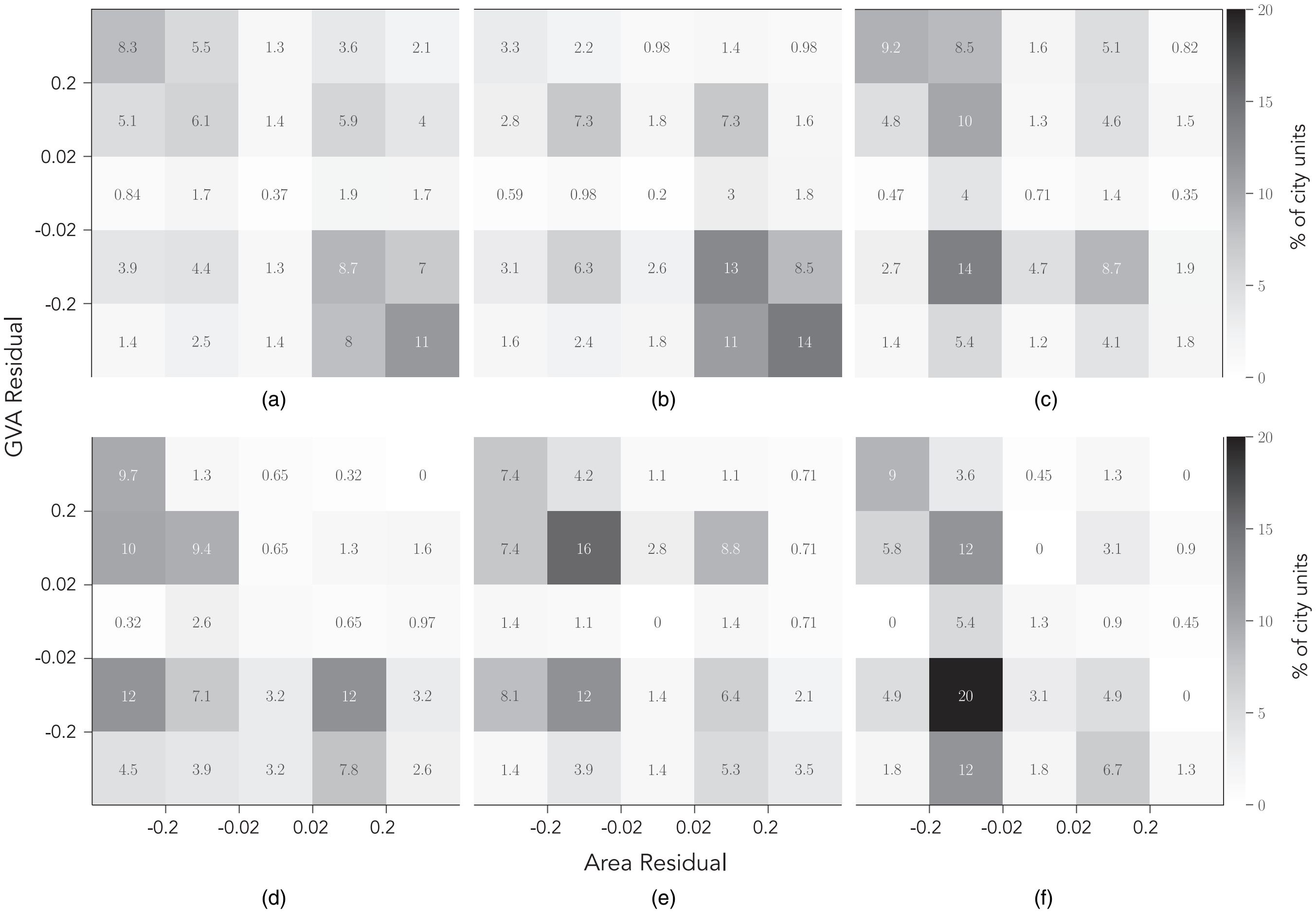


Fig. 6. Heat maps showing the percentage of city units across all boundaries for each residual cell: (a) DE; (b) NL; (c) EW; (d) Rhine–Ruhr; (e) Randstad; (f) Northern Powerhouse

are involved in the two variations of the Northern Powerhouse. The consideration of the scaling deviations, in contrast, highlights a pattern whereby the economic overperformance is virtually correlated with denser built areas. (It should be noted that the variation in the R^2 that is reported across panels in Fig. 5 is an artefact of the *modifiable areal unit problem* (Openshaw, 1983).) 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 to be borrowed from the two are perhaps not simply those concerning intercity mobility.

4.2. A subregional portrait of national differences

To complement the comparison of the three city regions and their home countries, we calculate the percentage of cities within different ranges of ξ_Y and ξ_{A_n}, building nationwide and regionwide city distributions. (Because of the overall similarity of city units and scaling regimes for the density-based boundaries, from this point forward, nationwide or regionwide aggregation of all units refers to all units within C100, C500, C1000, C1400, NUTS level 3, 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 the results presented in Fig. 6.) Fig. 6 shows discrete heat maps with residuals for urbanized area on the x-axis and that of economic output on the y-axis. The diagonal remains indicative of nearly optimal size–cost performance. The most noticeable difference between the nationwide distribution of city units in DE, NL and EW is the relative symmetry of the distribution about the diagonal in DE and NL mirroring their distributions in Fig. 3 with distribution peaks along the diagonal. Additionally, it is clear that these peaks in DE and NL are either units that are sparse and economically underperforming (the bottom right-hand quadrant) or those that are dense and economically overperforming (the top left-hand quadrant). This is in contrast with the EW 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 underperforming despite their perceived density (the bottom left-hand quadrant) with a size–cost balance in significant need of better internal mobility. Of more interest is the difference between national and regional distributions. Comparing the composition of the Randstad and Rhine–Ruhr with the overall German and Dutch distributions highlights a shift of the distribution peaks from sparse economically underperforming to denser and overperforming city units, especially in the Randstad, whereas a comparison of the Northern Powerhouse against the EW's composition reveals a slight increase in the portion of units that are both dense and underperforming.

5. Discussion

We can round up the findings of the analysis and our national and regional comparisons as

(a) continental case-studies, although very instructive, are not in themselves crucial in making a case for better transport infrastructure in England and Wales,
(b) better mobility is not the sole factor in the different agglomeration elasticities between DE, NL and EW, and,
(c) unlike the Randstad and Rhine–Ruhr, the Northern Powerhouse's performance is representative of the wider urban system framing the underperformance in EW as a national problem and not a purely local or regional problem.

Similarly to the results, we begin our brief discussion with the national comparison.

(a) *Continental case-studies, although very instructive, are not in themselves crucial in making a case for better transport infrastructure in England and Wales*. By using Bettencourt's systematic analytical framework which also enables comparisons of different urban systems, the Dutch and German case-studies can be seen as instructive for understanding and interpreting the UK evidence. 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 a systemwide 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. However, although the findings from the comparison between the three countries' urban networks are consistent with expectations, the Netherlands and Germany as national comparisons are found not to be crucial in arguing for better transport in England and Wales. The analysis of the UK data, in the light of the continental cases, is by itself shown to be sufficient to substantiate the case for mobility. In this manner, simply assessing EW's urban network in isolation with respect to the model's ideal could have supported a case for the deployment of better transport and mobility infrastructure, albeit those mostly of an intracity nature, for boosting the national economy and by extension that of the northern cities from an agglomeration point of view.

(b) *Better mobility is not the sole factor in the different agglomeration elasticities between DE, NL and EW*. 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 with the nearly linear scaling regime in England and Wales. This is in spite of a similarly linear scaling of urbanized area that is observed for both the Dutch and the 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 we might be able to extract transferable policy drivers from comparisons with better performing urban networks such as those of Germany and the Netherlands, a singularly intercity-transport-driven argument would not be the root solution or driver at which to arrive. A regional examination of the Rhine–Ruhr, the Randstad and Northern Powerhouse further reinforces this.

(c) *Unlike the Randstad and Rhine–Ruhr, the Northern Powerhouse performance is representative of the wider urban system*. We have shown that on average the Randstad and Rhine–Ruhr are comprised of individual units that themselves outperform individual units building up either realizations of their English counterparts economically. The consideration of the aggregated regions with respect to the scaling residuals appears to suggest that this shows an association with the higher densities of the continental examples demonstrated by the comparison of the three regions at different boundary definitions (Fig. 5), where the aggregated Northern Powerhouse shows considerably lower densities and by extension productivities. It is therefore notable that the only comparable economic overperformance of a Northern Powerhouse unit occurs at the 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 with a majority of units denser and overperforming in contrast with their national distributions. Meanwhile, the composition of the Northern Powerhouse is very much representative of England and Wales in general. This reframes the underperformance of the northern English units not as a regional problem but a problem at a national level. Nevertheless, the aggregate regional comparison, in contrast with the current transport-led infrastructural programme, suggests a need for further densification in the Northern Powerhouse by using the same agglomeration-based principles. On a related note, we have previously pointed to the difference between the geographic coverage of the *planned* Northern Powerhouse and its contiguously populated boundary; Fig. 4. Although insights from Fig. 5 suggest that this territorial difference does not influence size–cost optimality significantly, such geographic proximity issues could become influential when considering the practicality of implementing multiscale mobility improvements and/or densification measures.

Finally, an additional source of nuance, however, is the implication of singularly deploying either intercity 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 multiscale densification, i.e. shrinking area residual, whether or not the overall cost–size performance remains nearly optimal requires a balance between the two strategies to be reached. In this vein, Rhine–Ruhr can achieve higher potentials and size–cost balance through further improvements of mobility. The same is true for the Northern Powerhouse and the Randstad across a majority of spatial scales. Under the agglomeration economies paradigm, therefore, improvements and extensions of the intercity and intracity transport infrastructure become crucial not as the principle solution but as the complementary measures that are needed to maintain appropriate levels of mobility and hence the size–cost balance as any of the regions densify as a whole, across all or a given boundary definition, towards the top left-hand quadrant in Fig. 5.

6. Conclusions

The primary contribution of this paper rests in its use of urban scaling models to examine the applicability and transferability of intercity improvements to mobility in boosting economic productivity and output in the north of England. Our results show that, whereas intercity mobility and transport arguments can be used when considering overall national performance of urban networks, intercity transport solutions supported by stylized agglomeration-based arguments are not easily transferable from successful examples of polycentric metropolitan regions in boosting underperformance of similarly sized regions elsewhere. Indeed, when considering size–cost balance, an examination of the needs for better mobility and/or densification can be made without requiring external comparisons. This, at a first glance, may appear to paint such regional comparisons trivial. However, regional comparative approaches are essential in identifying certain nuances which cannot be identified by looking at single-case data. Indeed, this is the strength of Bettencourt's framework as it allows a parsimonious but sophisticated and coherent methodological framework to be applied in very different contexts. As such, although Bettencourt's framework demonstrates that the Dutch and German comparisons are not in themselves fundamental to the EW-specific arguments, this cannot have been known before the application of Bettencourt's framework. This additionally portrays the EW case simply as a specific example of a more general class of problems. The continental comparisons, in contrast, suggest that if mobility improvements do not drive and/or are not implemented in tandem with urban densification then these improvements are not likely to deliver the intended productivity gains on their own. This points towards a deeper interplay of productivity, population and infrastructural density. The paper then broadly argues in favour of mostly intracity transport and mobility infrastructure coupled with and supporting increased urban density in enhancing economic performance and productivity.

Acknowledgements

Maps of England and Wales contain Ordnance Survey data, Crown copyright and database rights 2018.

Maps of Germany and the Netherlands are copyright of EuroGeographics for the administrative boundaries.

References

Alstott, J., Bullmore, E. and Plenz, D. (2014) Powerlaw: a Python package for analysis of heavy-tailed distributions. *PLOS One*, **9**, article e85777.

Auerbach, F. (1913) Das Gesetz der Bevölkerungskonzentration. *Petrmnns Geog. Mitt.*, **59**, 74–76.

Batty, M. (2012) Building a science of cities. *Cities*, **29**, suppl. 1, S9–S16.

Bettencourt, L. (2013) The origins of scaling in cities. *Science*, **340**, 1438–1441.

Bettencourt, L. and Lobo, J. (2016) Urban scaling in Europe. *J. R. Soc. Interfc.*, **13**, article 20160005.

Bettencourt, L., Lobo, J., Helbing, D., Kühnert, C. and West, G. B. (2007) Growth, innovation, scaling, and the pace of life in cities. *Proc. Natn. Acad. Sci. USA*, **104**, 7301–7306.

Bettencourt, L., Lobo, J. and Youn, H. (2013) The hypothesis of urban scaling: formalization, implications and challenges. *Preprint arXiv:1301.5919.* Santa Fe Institute, Santa Fe. (Available from `http://arxiv.org/abs/1301.5919`.)

Bettencourt, L. and West, G. (2010) A unified theory of urban living. *Nature*, **467**, 912–913.

Burger, M. J., Meijers, E. J., Hoogerbrugge, M. M. and Tresserra, J. M. (2015) Borrowed size, agglomeration shadows and cultural amenities in North-West Europe. *Eur. Planng Stud.*, **23**, 1090–1109.

Centre for Cities (2015) Northern Powerhouse factsheet: key figures on the city regions in the Northern Powerhouse. Centres for Cities, London. (Available from `http://www.centreforcities.org/publication/northern-powerhouse-factsheet/`.)

Clauset, A., Shalizi, C. R. and Newman, M. E. J. (2009) Power-law distributions in empirical data. *SIAM Rev.*, **51**, 661–703.

Copernicus Land Monitoring Service (2016) Corine Land Cover 2012. European Environment Agency, Copenhagen. (Available from `https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012`.)

Dorling, D. (2010) Persistent North-South divides. In *The Economic Geography of the UK* (eds N. M. Coe and A. Jones), pp. 12–28. London: Sage.

Eurostat (2016a) GEOSTAT 2011 grid dataset. Eurostat, Luxembourg. (Available from `http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography/geostat`.)

Eurostat (2016b) Gross value-added at basic prices by NUTS 3 regions. Eurostat, Luxembourg. (Available from `http://ec.europa.eu/eurostat/web/products-datasets/-/nama_10r_3gva`.)

Eurostat (2017) Methodological manual on city statistics: 2017 edition. Eurostat, Luxembourg. (Available from `http://ec.europa.eu/eurostat/documents/3859598/8012444/KS-GQ-17-006-EN-N.pdf/a3f1004f-cfae-4cc4-87da-81d588d67ae2`.)

Gabaix, X. (1999) Zipf's law for cities: an explanation. *Q. J. Econ.*, **114**, 739–767.

Glaeser, E. L. (2010) Introduction to "Agglomeration Economics". In *Agglomeration Economics* (ed. E. L. Glaeser), pp. 1–14. Chicago: University of Chicago Press.

Glaeser, E. L. and Kohlhase, J. E. (2003) Cities, regions and the decline of transport costs. *Pap. Regnl Sci.*, **83**, 197–228.

Gomez-Lievano, A., Patterson-Lomba, O. and Hausmann, R. (2016) Explaining the prevalence, scaling and variance of urban phenomena. *Nat. Hum. Behav.*, **1**, article 0012.

Gomez-Lievano, A., Youn, H. and Bettencourt, L. (2012) The statistics of urban scaling and their connection to Zipf's law. *PLOS One*, **7**, article e40393.

Infrastructure and Projects Authority (2015) National infrastructure pipeline factsheet. *Technical Report*. Her Majesty's Treasury, London. (Available from `https://www.gov.uk/government/publications/national-infrastructure-pipeline-july-2015/national-infrastructure-pipeline-factsheet-july-2015`.)

Jones, C. (2016) Spatial economy and the geography of functional economic areas. *Environ. Planng* B, **44**, 486–503.

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

National Infrastructure Commission (2016) High speed North: a National Infrastructure Commission report. *Technical Report*. National Infrastructure Commission, London. (Available from `https://www.gov.uk/government/publications/high-speed-north-a-national-infrastructure-commission-report`.)

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

Organisation for Economic Co-operation and Development (2012) *Redefining "Urban": a New Way to Measure Metropolitan Areas*. Paris: Organisation for Economic Co-operation and Development.

Osborne, G. (2014) We need a Northern Powerhouse. (Available from `https://www.gov.uk/government/speeches/chancellor-we-need-a-northern-powerhouse`.)

Rozenfeld, H. D., Rybski, D., Gabaix, X. and Makse, H. A. (2011) The area and population of cities: new insights from a different perspective on cities. *Am. Econ. Rev.*, **101**, 2205–2225.

Samaniego, H. and Moses, M. E. (2008) Cities as organisms: allometric scaling of urban road networks. *J. Transprt Land Use*, **1**, 21–39.

Sim, A., Yaliraki, S. N., Barahona, M. and Stumpf, M. P. H. (2015) Great cities look small. *J. R. Soc. Interfc.*, **12**, article 20150315.

Smith, D. A. (2014) Domestic energy use in England and Wales: a 3D density grid approach. *Regnl Stud. Regnl Sci.*, **1**, 347–349.

Swinney, P. (2016) Building the Northern Powerhouse: lessons from the Rhine-Ruhr and Randstad. *Technical Report*. Centre for Cities, London. (Available from `http://www.centreforcities.org/wp-content/uploads/2016/06/16-05-31-Building-the-Northern-Powerhouse-Lessons-from-the-Rhine-Ruhr-and-Randstad.pdf`.)

Szüle, J., Kondor, D., Dobos, L., Csabai, I. and Vattay, G. (2014) Lost in the city: revisiting Milgram's experiment in the age of social networks. *PLOS One*, **9**, article e111973.

Transport for the North (2015) The Northern Powerhouse: one agenda, one economy, one North—a report on the northern transport strategy. *Technical Report*. Department for Transport and the Regions, London. (Available from `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427339/the-northern-powerhouse-tagged.pdf`.)

Transport for the North (2016) The northern transport strategy: spring 2016 report. *Technical Report*. Department for Transport and the Regions London. (Available from `https://www.gov.uk/government/publications/northern-transport-strategy-spring-2016`.)

Yakubo, K., Saijo, Y. and Korošak, D. (2014) Superlinear and sublinear urban scaling in geographical networks modeling cities. *Phys. Rev.* E, **90**, article 022803.

Supporting information

Additional 'supporting information' may be found in the on-line version of this article:

'Productivity, infrastructure, and urban density–an allometric comparison of three European city regions across scales'.