City boundaries and the universality of scaling laws

Elsa Arcaute1, Erez Hatna2, Peter Ferguson1, Hyejin Youn3, Anders Johansson4, and Michael Batty1

1Centre for Advanced Spatial Analysis (CASA), University College London, UK
2Center for Advanced Modeling, The Johns Hopkins University, USA
3Santa Fe Institute, USA
4Department of Civil Engineering, University of Bristol, UK

Keywords: Power-laws; Scaling laws; urban indicators; city boundaries

Abstract
This paper investigates the universality and robustness of scaling laws for urban systems, according to the work by Bettencourt, Lobo and West among others [7 10], using England and Wales as a case study. Initial results employing the demarcations for cities from the European Statistical Commission digress from the expected patterns. We therefore develop a method for producing multiple city definitions based on both morphological and functional characteristics, determined by population density and commuting to work journeys. For each of these realisations of cities, we construct urban attributes by aggregating high resolution census data. The approach produces a set of more than twenty thousand possible definitions of urban systems for England and Wales. We use these as a laboratory to explore the behaviour of the scaling exponent for each configuration. The analysis of a large set of urban indicators for the full range of system realisations shows that the scaling exponent is notably sensitive to boundary change, particularly for indicators that have a nonlinear relationship with population size. These findings highlight the crucial role of system description when attempting to identify patterns of behaviour across cities, and the need for consistency in defining boundaries if a theory of cities is to be devised.

Introduction
Every city has evolved under a unique set of geographical, political and cultural conditions [21] but despite heterogeneity in their historical trajectories, there appear to be certain characteristics common to all cities regardless of their location. Such characteristics include fractal properties [6 11], Zipf distributions of city sizes [35] and population growth laws [19 18 16 29 32]. In the search for appropriate paradigms, a single characteristic of urban systems, population city size, reveals scaling laws quantifying urban attributes ranging from innovation, income and employment rates, to household electrical consumption and road surface area amongst many others [7 9]. These are encompassed in the following relationship

\[ Y(t) = Y_0 N(t)^\beta \]  

where \( Y(t) \) and \( N(t) \) represent the urban indicator and the population size of a city at time \( t \) respectively, and \( Y_0 \) is a time dependent normalisation constant. Evidence of such laws has been observed in the US, Germany and China [10 8] among other countries. The exponent \( \beta \) has been found to lie within three universal categories [7]: (i) \( \beta < 1 \) (sublinear) economies of scale associated with infrastructure and services, e.g. road surface area; (ii) \( \beta \approx 1 \) (linear) associated with individual human needs, e.g. housing and household electrical consumption; and (iii) \( \beta > 1 \) (superlinear) associated with outcomes from social interactions, e.g. number of patents and income. A summary of the exponents found in [7 10] is given in Fig. 1A.

Scaling laws of this sort have proven to be very successful in biology [36 31], where size alone provides sufficient information to predict many properties of an animal, such as its life span, metabolism and heart rate. Consequently, identifying equivalent universal scaling laws for cities could be extremely important for furthering our understanding of urban dynamics, and may help to manage many contemporary global challenges that concern cities such as the effect of transport and industrial emissions on climate change, the use of natural resources and the growth of urban poverty [14].

Nevertheless, many challenges need to be overcome before universal laws can be established [29]. For example, questions remain as to how to generate reliable results both within and across national datasets. The problem of ambiguity in the definition and measurement of urban indicators gives rise to incongruent comparisons between countries. In addition, it is not always possible to specify the generative mechanism of an urban indicator, in order to unequivocally assign it as the outcome of either social interactions, human needs or services. Furthermore, the quality of the data might not permit a full characterisation of the exponent in either one of the three regimes: sublinear, linear and superlinear.

These challenges bring forward a more fundamental aspect in the development of a theory of cities. This refers to the robustness of the model, assessed through the sensitivity of the exponent to different city defini-
In this work, we analyse the scaling behaviour of a range of indicators using census data on cities in England and Wales. After initially adopting a predefined set of standard city delineations, we find that the observed scaling relationships depart from the expected behaviour. We hypothesise that the unanticipated results may be due to the given definition of city boundaries. In response, we explore new methodologies to generate more realistic city boundaries from more disaggregate data. Our method identifies city extent by clustering very small scale geographic units according to population density and journey to work commuting trips, removing the need for an a-priori assumption about boundary demarcation. This enables us to define cities based on both their morphological and functional extent. A series of urban areas is generated by clustering adjacent neighbourhoods that lie within a defined density threshold. Then the effective commuting hinterland of each core is identified by calculating the proportion of journey to work trips that are destined for each core. We can then probe the whole parameter space of density and commuting thresholds to obtain more than $20 \times 10^3$ realisations of cities. When the scaling laws are tested for all these system descriptions, we find that the exponent is consistent for variables in the linear regime, but that variables in the non-linear regimes are highly sensitive to boundary definition and sample distribution. These results emphasise the need for a consistent methodology to define urban systems, in order to derive congruent comparisons between cities. They also reveal the importance of understanding the origin of deviations from a model, in order to construct a theory of cities.

### Results

**Scaling laws in England and Wales (E&W)**

The complex, multi-layered nature of cities means that more than one reasonable definition of their extent can be produced depending on whether the political, economic or geographic reach of the system is being considered \[^3,4\]. Although this plurality is a well known problem in the study of urban systems, there is no consensus as to how a city should be defined. For the purposes of an initial exploration of scaling behaviour in England and Wales (E&W), we ignore the challenge of boundary delineation by adopting the definition employed in \[^7\], where cities are considered as integrated economic and social units. In the case of the US this definition corresponds to Metropolitan Statistical Areas (MSAs)\[^5\], and an analogous set of areas for the European Union are Larger Urban Zones (LUZs)\[^1\]. We select a range of observed socio-demographic indicators produced by the UK Office for National Statistics (ONS) from the 2001 census for E&W, and aggregate high resolution census units up to the LUZ classification. We exclude Scotland from the analysis since the National Records of Scotland applies a different methodology for data collation than the ONS in England and Wales (details of all data sources are provided in the SI Text in the ‘Individual Data Tables’ section). Analysis of scaling laws is then undertaken by classifying each indicator in terms of the above mentioned three regimes: sublinear, linear and superlinear.

When attempting to classify the resulting scaling behaviour from this initial exploration, many of the indicators could not be placed within a unique domain. See Table S1 and discussion in SI Text. Notwithstanding, there are some urban indicators in the list that can be clearly categorised in one of the three regimes: sublinear, linear and superlinear.

In Fig. 1B, where \(\beta\) is the exponent in eq. (1). These exponents can be found in Table S2.

Although we observe many variables that do behave as expected in the sublinear and linear regime, given the confidence intervals associated with each exponent, further verification is clearly required for some of them. On the other hand, the values for the scaling exponent in the superlinear regime do not corroborate the expected ones. This is particularly surprising for three variables that are clearly outcomes of social and economic interactions: number of patents, household income and crime incidents. The latter can be regarded as one kind of social activity, that therefore is expected to increase su-
perlinearly with city size as discussed in [8, 20].

Let us now investigate if all these discrepancies are due to an inappropriate definition of cities. Fig. S2 gives a representation of cities in E&W in terms of LUZs and their size distribution. The map shows a high degree of arbitrariness in the selection of cities and the delimitation of their boundaries. The size distribution is more or less Zipf. However, the next biggest cities after London seem to be underestimated according to this plot. The misrepresentation of these cities out of a total of 21, in addition to the absence of important cities such as Oxford and Reading, questions the soundness of the LUZ representation.

In the next section we look into new ways of elucidating cities in E&W in an attempt to answer the question about the influence of boundary definition on observed scaling behaviour. Our first aim is to look for contours of cities that are consistent with the built environment and their economic flows. We achieve this by using population density as a fundamental urban property to construct the initial settlements. This is followed by an expansion of the boundaries driven by commuting to work flows. Our second aim is to explore scaling laws in a comprehensive set of systems of cities, so that we can analyse the behaviour of the scaling exponent under these different definitions. This will give us insight into the stability or sensitivity of such laws to city boundaries.

Redefining city boundaries through density

In order to redefine cities, we construct a clustering algorithm parametrised by population density. A similar algorithm can be found in [30]. The unit of agglomeration is a ward, which is the smallest geographical unit in the census data across many variables (see SI Text, ‘Unit of Geography’ section for details). Let $\rho$ be our density parameter. We cluster all adjacent wards with density $\rho_w < \rho$, but is surrounded by wards such that $\rho_w \geq \rho$, then the ward is also included in the cluster. This is done in order to avoid cities with holes. For example, if a ward contains a big park, such as in Richmond in Greater London, its density will be much lower than its adjacent wards. If left out of the cluster, the city will not only have a hole, but will be missing an important functional area. Cities are hence considered as continuous entities in this first approach. In detail, the parameter $\rho$ is varied within the interval $[1; 40]$ persons/hectare. The result is 40 different realisations of systems of cities for E&W, varying from very large clusters containing various settlements, to clusters containing only the core of cities for the highest density values. Maps for some of these cases are featured in the SI Text (see Fig. S3). When we follow the growth of cluster sizes resulting from the change in density from high to low values, we observe a sharp transition of the rank 3 cluster between $\rho=13$ and $\rho=12$, see Fig. 2A. This transition corresponds to the joining of Liverpool and Manchester. The biggest cluster encompasses London, and this grows steadily including small settlements as the density lowers, but does not merge with another big city within the interval considered, and this is why no transition is observed. If we select a density threshold before the merging of Liverpool and Manchester takes place, but near the transition since these two cities are very close spatially, we reproduce a system of cities that is very similar to the one prescribed by the built environment. This corresponds to $\rho = 14$, and we see from Fig. 2B that it recreates almost exactly the urbanised areas defined using the CORINE landcover dataset [15]. In addition cities follow a Zipf distribution. Fig. 2B shows the cumulative density function with an exponent of 2.07 and a very high p-value of 0.8, using the method for fitting a power-law distribution proposed in [12].

Extending boundaries to include commuters

At this point, we expand our definition of the city from a purely morphological description to include some sense of its functional extent through an incorporation of commuting flow data into the clustering algorithm. Data on the total journey to work flows of commuters from every ward in E&W is provided by the ONS and is used to define the commuting hinterland for the original 40 cluster systems. The procedure operates as follows. For each realisation for $\rho\in[1; 40]$, we select only clusters whose population size $N$ is such that $N \geq N_0$, where $N_0 \in \{0, 10, 50, 100, 150\} \times 10^3$ individuals. Each ward is then added to the cluster for which the largest percentage $\tau$ of people commute into if $\tau > \tau_0$, where $\tau_0 \in [0; 100]$. No continuity condition is imposed on the new clusters. The extreme value of $\tau_0 = 100$ reproduces the original system. This procedure leads to a comprehensive list of $20.2 \times 10^3$ realisations of systems of cities. See the SI Text for a visual representation of some of these composites (Fig. S4). Exploring the full parameter space is useful to assess the behaviour of the scaling exponent, bearing in mind that for the extreme values of $\rho$ and low values of $\tau$, the sets of aggregates move further and further away from realistic descriptions of cities.

Sensitivity of the power law exponent

We make use of heatmaps to represent the values of $\beta$ in eq. (1), for each of the five initial population cut-offs pre-commuting clustering: $N \geq N_0$, in the whole
parameter space for $\tau_0 \in [0; 100]$ and $\rho \in [1; 40]$. Linear relationships of variables with population size, are generally not affected by the different definitions of cities presented above. For this reason, heatmaps for these variables tend to be homogeneous over the whole parameter space. On the other hand, non-linear dependencies coming from aggregation effects, will exhibit variations in the scaling exponent if initial conditions, given by the city limits and population aggregates through commuting patterns, are changed.

A first inspection indicates high variability between heatmaps for the same urban indicator but different population size cutoffs. The effect of imposing a minimum size on a settlement in order to consider it a city, is the reduction in the number of cities that are included in the system to test for scaling laws, see Fig. S4. For the extreme scenarios, i.e. very high density and minimum population size of $150 \times 10^3$, the number of cities included in the distribution can vary greatly, from 429 with no cutoff, to only 5 cities if a large population size cutoff is imposed. Variations between the different heatmaps for the extreme values of the density parameter, are therefore mainly due to the small sample size in the distribution, and no statistically sound conclusion can really be drawn from these extreme cases. This is particularly the case for Income as shown in Fig. S3. For the extreme scenarios, i.e. where there is only a very small number of cities included in the distribution, the weight of London becomes important, biasing the exponent to superlinearity. London is a positive outlier, and relevant economic agglomeration effects are clearly present, but do not seem to exist for the next biggest cities. We also observe that in this case no effects are recorded by including commuters into the analysis. For properties that belong to the linear regime, and where London is not an outlier, such as number of households and of people employed, the exponent remains consistent across the whole parameter space, see Fig. 4.

Conversely, the exponent for the observables in the sublinear regime can be greatly affected by variations in the threshold for the percentage of commuters. See Fig. 5 for employment in agriculture. In this case, extending the area of cities to include commuters, makes the exponent shift from the sublinear to the superlinear regime. This effect is pronounced if in addition settlements are removed through the constraint of minimum population size of $10^5$. Furthermore, for the range where clusters are just the central cores of cities, given by $\rho > 30$, the exponent also increases if population size constraints are applied. For variables such as distance to work, see Fig. S6, the exponent becomes superlinear, although as stated earlier, for these extreme cases no statistical validity can be obtained. There are also many variables that should present economies of scale, such as infrastructure variables, e.g. area of roads; and some employment categories such as elementary occupations and manufacturing, that nevertheless show a linear exponent if no cutoff on population size is imposed, see Fig. S7. If on the other hand only the tail of the distribution containing the large cities is taken into account, the exponent tends to sublinearity, although very weakly, and only for some of these variables.

The expected agglomeration effects for variables that are the outcome of social and economic interactions are in general not observed. Most of the employment categories corresponding to this regime have linear

Figure 3: Heatmap for income for different minimum population size thresholds: no cutoff, $10^4$, $10^5$ and $150 \times 10^3$.

Figure 4: Heatmap for observables that belong to the linear regime. The exponent remains consistent in the whole parameter space.

Figure 5: Heatmap for employment in agriculture, hunting and fishing for different minimum population size thresholds: no cutoff, $10^4$, $10^5$ and $150 \times 10^3$.
...exponents, see Fig. S8 with the exception of Financial Intermediation. Nevertheless the value of the exponent is nowhere close to the expected 1.15 from supereconomic employment or 1.30 from R&D employment. In addition, once the cities are extended as integrated economic entities by including commuters, the effect on the exponent is to lower its value towards linearity, instead of increasing it. The heatmaps show how these non-linear effects are highly sensitive to city definition and sample distribution, see Fig. 6. Different cutoffs on population size give very different results. And once again, the effect of London as a positive outlier becomes important for a small sample size. Other variables displaying superlinearity are area of non-domestic buildings and patents. For the former the superlinearity, which is considerable if no constraint on population size is imposed, is completely washed out if a minimum population size of $50 \times 10^3$ individuals is imposed, see Fig. 7. Patents on the other hand, present the higher volatility, and each heatmap for different population size cutoffs gives a very different result, see Fig. S8. For this variable we need to impose a minimum population size threshold of $10^4$ in order to obtain a dataset that does not contain many zeroes from the small settlements.

In conclusion, our results show that cities in E&W do not present economic agglomeration effects, with the exception of London, which is a positive outlier. Furthermore, if non-linear effects are present, the scaling exponent is highly sensitive to city delimitation. Therefore the value of the exponent for a single definition cannot be taken as a proxy to draw comparisons between cities if no consistent way to construct cities has been devised. On the other hand, we were also able to confirm that all the urban indicators that have a linear dependency with population size, are robust to city demarcations. Any discrepancies found between observed linear dependencies and expected non-linear ones according to 7, 8, are therefore not due to a poor definition of cities, since more than $20 \times 10^3$ different configurations were explored, and linearity persisted. The main incongruity between observed and expected outcomes, is the lack of superlinearity for exponents belonging to observables that are the product of economic and social dynamics, such as income and some employment categories requiring particular skills.

Our methodology provides a tool to construct multiple city limits in a systematic way, enabling us to define consistent systems of cities. Moreover, it emphasises the sensitivity to boundary definition for urban indicators that do not show a linear dependency with population size. This is specifically highlighted for patents, whose exponent is highly volatile across the whole parameter space. The sensitivity of non-linear exponents to boundaries indicates that comparisons drawn between cities based on the value of the exponent can be misleading.

**Discussion**

This work shows that the search for patterns in urban indicators is more intricate than previously thought. The specific demarcation and definition of cities play a crucial role in the distribution and measurement of urban attributes. Any dependencies found between the latter and population size, are strongly affected by city selection and definition. The argument could however be reversed if universality existed. Contours of cities could be constructed according to expected statistical or scaling laws. This can only be done if these are not too sensitive to borders, since otherwise one would face the dilemma of what comes first: the boundary upon which the theory rests, or the theory that defines the boundary.

The lack of superlinear exponent values for indicators driven by social and economic interactions sets a significant discrepancy between results for E&W, and results found for the US and China in the literature. This raises many important questions. On the one hand, the distinct results obtained from different population size cutoffs, indicate that maybe E&W is too small a system from which agglomeration effects can be measured prop-

---

**Figure 6:** Heatmap for employment in financial intermediation for different minimum population size thresholds: no cutoff, $10^4$, $10^5$ and $150 \times 10^3$.

**Figure 7:** Heatmap for area of non-domestic buildings for different minimum population size thresholds: no cutoff, $10^4$, $10^5$ and $150 \times 10^3$. 
erly. This also prompts us to investigate the necessary conditions that need to be fulfilled in order to observe the expected scaling exponents. In addition to the relatively diminutive scale of England and Wales compared to the US and China, the spatial distribution of its cities might be responsible for obscuring superlinear dependencies. Individual cities within the UK are perhaps too close to each other to be easily treated as discrete entities, or maybe London is simply too big relative to the rest of the system. This can be investigated further, by looking for agglomeration effects in other countries that contain a primate city such as London [22].

Alternatively, the main discrepancies might be due to the fact that the UK has experienced a process of de-industrialisation for longer than most western countries. The combined impact of globalisation and de-industrialisation may be causing a slow down in the growth of the largest cities outside of London, in turn affecting the value of the exponent. This affect might be particularly noticeable in the results, as the exponent of a power law is driven by the largest events in the distribution and the weight of London is not big enough to skew the exponent. This hypothesis can also be put to the test by calculating the exponent for various indicators at the height of the British industrial revolution, during the growth phase of the above mentioned industrial cities. The scaling behaviour during this time period could be more in line with the expected scaling laws drafted in [7 8]. If this is the case, perhaps there is a timeline that all countries follow after industrialisation that alters expected scaling behaviour?

On the other hand, the maturity of the UK with respect to industrialisation, makes it a unique integrated urban system [5], in which, following governmental policies of regionalisation and decentralisation [11], critical functionalities were removed from the core of main cities and placed in other areas, smoothing away any agglomeration effects from economic output. Such regionalisation policies were applied by the government from the 1920s onwards, with respect to reducing congestion of employment in London, stemming the so called 'drift to the south' and attempting to reinvigorate all regions outside London and south east England.

The geographic scale of the system of interest might need to reflect the reach of a city’s interactions. London’s dominance as a financial and business services hub relates as much to a global organisation of trade and interaction. A characteristic which is reflected in the regressions as London is a positive outlier for attributes that are expected to have superlinear dependencies. The performance of cities such as London should possibly be evaluated relative to other global hubs operating within a larger scaled network of interactions. Following Sornette’s idea on the emergence of big things [33, 37, 28], a different perspective of the description of cities could be adopted, in which these global hubs are evaluated separately to their domestic counterparts. Sornette refers to the formers as dragon-kings. A two system theory of cities might then emerge. A regime for cities driving international dynamics, the dragon-kings, and a regime for the remaining cities composing a country.

The methodology employed in this paper, had the purpose of recreating several representations of cities in England and Wales in order to explore the sensitivity of the scaling exponent, and to distinguish between linear and non-linear effects, by looking at the fluctuations between different realisations. Across more than twenty thousand descriptions, we were able to record and characterise the behaviour of the scaling exponent over the whole set. It is our intention to apply this methodology to other European countries, distinguishing between systems with and without primate cities. In addition, we intend to re-analyse the US by applying clustering algorithms to more disaggregate data than MSAs, and to assess historic datasets for the UK to evaluate the stability of the scaling exponent over time as well as space. The challenge of studying cities in a consistent manner is clearly considerable due to spatial and qualitative differences in data from location to location. It is however a necessary step for identifying truly universal patterns of behaviour. It is our hope that the approach described here goes some way to meeting this challenge, progresses the debate on scaling and city size and moves us closer to a better understanding of systems of cities.

Materials and Methods
Most of the variables come from the 2001 UK census dataset, produced by the Office for National Statistics. The data is of high spatial resolution, and is given at the level of wards. This is aggregated for each of the different realisations of cities described in the text. Each of the tables from which the indicators were obtained is described in detail in the SI Text. Data on patents was provided by the intellectual property office at the postcode level, for the years 2000 to 2011. The dataset on household income was taken from UK census experimental statistics for 2001/02, and it was produced using a model-based process. Crime data was obtained from the Home Office, and is the average between the years 2003 and 2011. Finally, infrastructure data, such as the area of roads, paths and buildings, comes from the 2001 Generalised Land Use Database. Details for all the variables are provided in the SI Text.
Acknowledgements

EA, EH, PF, AJ and MB acknowledge the support of ERC Grant 249393-ERC-2009-AdG. Useful discussions with Luís Bettencourt, and Geoffrey West of the Santa Fe Institute, and José Lobo of Arizona State University helped clarify many issues. HY acknowledge the support of grants from the Rockefeller Foundation and the James McDonnell Foundation (no. 220020195).

References

[1] M. Batty. The size, scale and shape of cities. Science, 319:769, 2008.
[2] M. Batty. Visualizing spacetime dynamics in scaling systems. Complexity, 16(2):51–63, 2010.
[3] M. Batty. Cities, prosperity, and the importance of being large. Environment and Planning B: Planning and Design, 38(3):385–387, 2011.
[4] M. Batty. Commentary: When all the worlds a city. Environment and Planning A, 43(4):765–772, 2011.
[5] M. Batty and P. Ferguson. Defining city size. Environment and Planning B: Planning and Design, 38(5):753–756, 2011.
[6] M. Batty and P. Longley. Fractal Cities: A Geometry of Form and Function. Academic Press, San Diego, CA and London, 1994.
[7] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West. Growth, innovation, scaling, and the pace of life in cities. Proc. Natl. Acad. Sci. USA, 104(17):7301–7306, 2007.
[8] L. M. A. Bettencourt, J. Lobo, D. Strumsky, and G. B. West. Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. PLoS ONE, 5(11):e13541, 2010.
[9] L. M. A. Bettencourt and G. B. West. A unified theory of urban living. Nature, 467:912–913, 2010.
[10] L. M. A. Bettencourt, J. Lobo, and G. B. West. Why are large cities faster? Universal scaling and self-similarity in urban organization and dynamics. Eur. Phys. J. B, 63(3):285–293, 2008.
[11] T. Champion, M. Coombes, S. Raybould, and C. Wymer. Migration and Socioeconomic Change: A 2001 Census Analysis of Britain’s Larger Cities. Policy Press, London, 2007.
[12] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. SIAM Review, 51(4):661–703, 2009.
[13] M. Cristelli, M. Batty, and L. Pietronero. There is more than a power law in Zipf. Sci. Rep., 2:812, 2012.
[14] M. Davis. Planet of Slums. Verso, London, 2006.
[15] EEA. Corine land cover update 2000, technical guidelines. European Environment Agency. ISBN: 92-9167-511-3, Copenhagen, 2002.
[16] J. Eeckhout. Gibrat’s law for (All) cities. Am. Econ. Rev., 94(5):1429–1451, 2004.
[17] EUROSTAT. Urban Audit, Methodological Handbook. Office for official Publications of the European Communities. ISBN 92-894-7079-8, Luxembourg, 2004.
[18] X. Gabaix. Zipf’s law and the growth of cities. The American Economic Review, 89(2):129–132, 1999.
[19] X. Gabaix. Zipf’s law for cities: An explanation. The Quarterly Journal of Economics, 114(3):739–767, 1999.
[20] A. Gomez-Lievano, H. Youn, and L. M. A. Bettencourt. The statistics of urban scaling and their connection to Zipf’s law. PLoS ONE, 7(7):e40393, 2012.
[21] P. Hall. Cities in Civilisation: Culture, Innovation and Urban Order. Weidenfeld and Nicholson, London, 1998.
[22] M. Jefferson. The law of the primate city. Geographical Review, 29:226–232, 1939.
[23] B. Jiang and T. Jia. Zipf’s law for all the natural cities in the united states: a geospatial perspective. Int. J. of Geo. Inf. Sci., 25(8):1269–1281, 2011.
[24] M. E. J. Newman. Power laws, pareto distributions and Zipf’s law. Contemp. Phys., 46(5):323–351, 2005.
[25] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo. A tale of many cities: Universal patterns in human urban mobility. PLoS ONE, 7(5):e37027, 2012.
[26] R. Perline. Weak and false inverse power laws. Statistical Science, 20(1):68–88, 2005.
[27] C. M. A. Pinto, A. Mendes Lopes, and J.A. Tenreiro Machado. A review of power laws in real life phenomena. Commun Nonlinear Sci Numer Simulat, 17:3558–3578, 2012.
[28] V. F. Pisarenko and D. Sornette. Robust statistical tests of dragon-kings beyond power law distributions. European Physical Journal, Special Topics (special issue on power laws and dragon-kings), 205:95–115, 2012.
[29] H. D. Rozenfeld, D. Rybski, J. S. Andrade, Jr., M. Batty, H. E. Stanley, and H. A. Makse. Laws of population growth. Proc. Natl. Acad. Sci. USA, 105(48):18702–18707, DEC 2 2008.
[30] H. D. Rozenfeld, D. Rybski, X. Gabaix, and H. A. Makse. The area and population of cities: new insights from a different perspective on cities. Am. Econ. Rev., 101:2205–2225, 2011.
[31] K. Schmidt-Nielsen. Scaling in Biology - Consequences of size. J. Exp. Zool., 194(1):287–308, 1975.
[32] F. Sembolini. Hierarchy, cities size distribution and Zipf’s law. Eur. Phys. J. B., 63:295–301, 2008.
[33] D. Sornette and G. Ouillon. Dragon-kings: mechanisms, statistical methods and empirical evidence. European Physical Journal, Special Topics (special issue on power laws and dragon-kings), 205:1–26, 2012.
[34] M. P. H. Stumpf and M. A. Porter. Critical truths about power laws. Science, 335:665–666, 2012.
[35] USC Bureau of Economic Analysis. United States Census Bureau. http://www.census.gov/geo/reference/garm.html, 1994.
[36] G. B. West, J. H. Brown, and B. J. Enquist. A general model for the origin of allometric scaling laws in biology. Science, 276(122), 1997.
[37] V. I. Yukalov and D. Sornette. Statistical outliers and dragon-kings as bose-condensed droplets. European Physical Journal, Special Topics (special issue on power laws and dragon-kings), 205:53–64, 2012.
[38] G. K. Zipf. Human Behavior and the Principle of Least Effort. Addison-Wesley, Cambridge, MA, 1949.
Supporting Information

**Abbreviations:** CAS, Census Area Statistics; ST, Standard Table; SIC, Standard Industrial Classification

**SI Text**

**Unit of Geography**
The underlying spatial unit for all city cluster aggregations is the Census Area Statistics (CAS) ward definition produced by the UK Office for National Statistics. Ward boundaries reflect the political geography of the UK at a fine resolution and due to the need to maintain equality of representation in political elections, have similar populations. CAS (Census Area Statistics) ward boundaries in particular have been the standard format for the release of ward level census information since 2003. They reflect electoral ward boundaries promulgated as at 31/12/2002 and contain 8850 separate wards for England and Wales.

Much of the 2001 census data was initially aggregated into a previous definition of ward boundaries known as Standard Table (ST) wards which are closer to the original electoral ward boundaries. These original ward boundaries contain 18 wards with fewer than 100 residents or 40 households so these small wards have been merged with other wards to protect data confidentiality and create the CAS ward definition. Further information on ward boundary definitions and for the conversion table between the older ST ward definition and the contemporary CAS ward definition can be found at [1, 2].

General information on census geography including information on the geographic hierarchy of the various units including Lower Layer Super Output Areas (LSOAs), Middle Layer Super Output Areas (MSOAs), Wards, Local Authorities (LAs), Government Office Region (GOR) and geometric centroid definitions can be found at [1, 2].

Census data used in the study was only provided for England and Wales as the process for collating data in Scotland and the definition of geographic boundaries meant that equivalent datasets could not be produced. More information on census geography for Scotland for the 2001 census can be found at [3].

**Individual Data Tables**

The original data and associated metadata for tables UV02, UV53, KS15, UV34, KS12 and Income can be found under the topics section of the UK neighbourhood statistics website [1].

**Population (Table UV02)**
The data on population was taken from UK census table UV02. Population data was taken from a data table on population density at the CAS ward level which provided separate statistics for total population, ward area and a result in population density figure. The total population figure was used for all regressions with socio-demographic variables used in the study.

**Housing Stock (Table UV53)**
Data on household dwelling numbers comes from census table UV53. The table provides information on the number of households, occupied or unoccupied, within each ward. Unoccupied household spaces are split into second residences/holiday accommodation, and vacant household spaces. A household space is the accommodation occupied by an individual household or, if unoccupied, available for an individual household. The population of this table is therefore all household spaces. The category used for regression was all household spaces and therefore included all spaces whether they were occupied or unoccupied.

**Travel to Work (Table KS15)**
Data on travel to work distances was taken from the UK census table KS15. The table shows both the length and the means of travel to work used for the longest part, by distance, of the usual journey to work. For the purposes of this table, public transport is defined as Underground, metro, light rail or tram, train and bus, minibus or coach. The distance travelled to work is the distance in kilometres of a straight line between the residence postcode and workplace postcode. The distance is not calculated for people working mainly at or from home, people with no fixed workplace, people working on an offshore installation or people working outside the UK. The population of the table is all people aged 16 to 74 in employment.

**Industry of Employment (Table UV34)**
Data on the industry of employment of resident employees was taken from UK census table UV34. The table shows the usual resident population aged 16 to 74 in employment by the industry they work in. The industry in which a person works is determined by the response to the 2001 census question asking for a description of the business of the persons employer (or own business if self-employed). The responses were coded to a modified version of the UK Standard Industrial Classification of Economic Activities 1992 UK SIC (92). The population in each category calculated for all people aged 16 to 74 in employment.

In the 2001 census, industry of employment information was collected for usual residents. A usual resident was generally defined as someone who spent most of their time at a specific address. It included: people who usually lived at that address but were temporarily away (on holiday, visiting friends or relatives, or temporarily in a hospital or similar establishment); people who worked away from home for part of the time; students, if it was their term-time address; a baby born before 30 April 2001 even if it was still in hospital; and people present on census day, even if temporarily, who had no other usual address. However, it did not include anyone present on census day who had another usual address or anyone who had been living or intended to live in a special establishment, such as a residential home, nursing home or hospital, for six months or more.

The category name for industry of employment used in the study were the following: Agriculture, hunting and forestry; Manufacturing; Construction; Hotels and Restaurants; Financial Intermediation; Real Estate, renting and business activities; Public administration, defence and social security; Education.

**Occupational Groups (Table KS12a)**
Data on occupational groups was taken from UK census table KS12. The information on this table comes from responses to questions asking for the full title of the main
job and description of what is done in that job from the 2001 census. The population of the table is all people aged 16 to 74 in employment and the values see are the absolute count values as opposed to the percentage values also provided.

For employment related data, any person who carried out paid work in the week before the census, whether self-employed or an employee, is described as employed or in employment. 'Paid work' includes casual or temporary work, even if only for one hour; being on a government-sponsored training scheme; being away from a job/business ill, on maternity leave, on holiday or temporarily laid off; or doing paid or unpaid work for their own or family business. A person’s occupation is coded from the responses to the questions asking for the full title of the main job (the job in which a person usually works the most hours).

Responses are coded to the Standard Occupational Classification 2000 (SOC 2000). Where possible census results are presented using standard classifications. The category used for regression were the following: Managers and Senior Officials; Professional occupations; Associate Professional and Technical operations; Skilled trades occupations; Administrative and Secretarial Occupations; Personal Service Occupations; Sales and customer service occupations; Process; plant and machine operatives; Elementary Occupations (examples of elementary occupations include Farm Workers, Labourers, Kitchen Assistants and Bar Staff).

Patents
Patent information was provided by the intellectual property office with postcode level reference that was subsequently aggregated to the CAS ward level. Data was provided for the years 2000 to 2011 inclusive to ensure sufficient quantity to avoid null values for individual wards. The total number of patents in the dataset that could be identified in E&W was 66,270. The values used for regression were simply the gross number of patents registered to a particular postcode whether it be business or home address.

More information on patent information from the UK IPO can be found at [4].

Household Income
The dataset on household income was taken from UK census experimental statistics for 2001/02 and is provided at a fine geographic resolution for the whole of England and Wales. The original data and associated metadata for Household Income can be found under the topics section of the UK neighbourhood statistics website [1].

The income data was produced using a model-based process which involves finding a relationship between survey data (data available on income) and other data drawn from administrative and census data sources. A model fitting process is used to select co-variates with a consistently strong relationship to the survey data. The strength of the relationship with these covariates is used to provide estimates on income for those wards where survey data on income is not available. More information on the provenance of the income data can be found on the appropriate page of the UK neighbourhood statistics census access site.

The survey data on income was taken from the Family Resources Survey (FRS) for the same year (2001/02). The total sample size for the 2001 survey was 42,000 addresses taken from across the UK. The FRS provides four variables that can then be generated for the whole country:

1) Average weekly household total income (unequalised).
2) Average weekly household net income (unequalised).
3) Average weekly household net income before housing costs (equalised by McClements equivalence scale).
4) Average weekly household net income after housing costs (equalised by McClements equivalence scale)².

Total income gross earnings, investments, pension provisions and welfare payments is the closest representation of the gross earning power of a given ward. This is the income variable associated with the regressions discussed in the results section.

Crime Data
The data on annual incidence of crime was obtained from the Home Office web site [7]. This data consists of the number of occurrences of different categories of crime (excluding homicides) at the local authority level annually between 31st March 2003 to 31st March 2011. The data analysed is the average number over the period, covering England and Wales.

Land Use Statistics
The Generalised Land Use Database (GLUD) figures show the areas of different land types for census Output Areas (OAs), Lower Layer Super Output Areas (LSOAs), Middle Layer Super Output Areas (MSOAs), Local Authorities (LAs), and Government Office Regions (GORs) in England as at 1st November 2001. Output level data was aggregated to the ward level for comparative analysis with population.

For the GLUD, a classification has been developed which allocates all identifiable land features on the UK Ordnance Survey MasterMap national mapping product into nine simplified land categories and an additional ‘unclassified’ category. These are:

1. Domestic buildings;
2. Non-domestic buildings;
3. Roads;
4. Paths;

¹The FRS is produced by the UK Department for Work and Pensions (DWP) to ensure a large sample sizes when collating information on household expenditure. Information on the FRS for 2001 can be found on the research section of the dwp.gov website [9], the methodology section (section 8) of the FRS summary report for that year at [9] and the associated technical report available through the ONS [2].
²The McClements equivalence scale adjusts income according to the relative advantage or disadvantage associated with households of different sizes, primarily to take into account the effect of economies of scale on household budgets of household size.
5. Rail;
6. Gardens (domestic);
7. Greenspace;
8. Water;
9. Other land uses (largely hardstanding); and
10. Unclassified.

The statistics are created by identifying different land parcels and buildings on an Ordnance Survey digital map product, and records their type and area. Each land parcel is then assigned to a specific Output Area based on its central point, and the information is aggregated to higher geographies. The building blocks for the statistics are a combination of objects in the electronic Ordnance Survey MasterMap product and AddressPoint business information (see below). The combination of MasterMap attributes, contextual analysis, and AddressPoint(TM) information provides the basis for the nine generalised land classes. Each polygon on MasterMap has attributes associated with it, and these provide information about the type of land covered, which can be used as a basis for generating a land use classification.

Ordnance Survey MasterMap is a large scale digital map for use in geographical information systems (GIS) and database systems. Real world objects are represented as explicit features, by polygons, each identified by a unique number called the TOID (topographic identi- fier) Each polygon also has digital attributes such as a theme, which enable querying and searching for specified features. The TOID enables linking to other datasets, including Ordnance Survey Address-Point(TM), via in this case, a linkable dataset called the National Build- ings Data Set (NBDS). This release of Generalised Land Use Database statistics uses the original November 2001 version of MasterMap.

Ordnance Survey Address-Point is a dataset that uniquely defines and locates residential, business and public postal addresses. It is created by matching information from OS digital map databases with more than 25 million addresses recorded in the Royal Mail Postcode Address File (PAF). The Generalised Land Use Database for England uses both Address-Point(TM) (version 2002.2.1, Dated 27 May 2002) and the National Buildings Dataset (NBDS) (July 2002). The NBDS is used as a link dataset to match Address-Point records to MasterMap records. All building TOIDs were classified as ‘Domestic Buildings’, unless any one or more of the following conditions were met:

a) it was seen to be adjacent to an area of hard-standing (such as a tarred car park or estate road) which was more than 300 square metres; b) it contained an address point with a business or organisation name; or c) it had an area greater than 1,000 square metres and did not contain any address point. Building TOIDs fulfilling any one or more of these criteria were recorded and classed as ‘Non-Domestic Buildings’. The category ‘Other’ is largely areas of hardstanding such as car parks, estate roads and hard tennis courts.

Urban indicators for Larger Urban Zones in E&W

We point out that that the classification of urban indi- cators in terms of three unique categories is not always possible. For example, one would expect that highly skilled employment categories would lie in the superlinear regime and conversely, basic employment categories would be expected to lie in the sublinear regime. However, it is clear from Table S1 that this is not the case for the LUZ definition. Two very similar employment categories, i) public administration and defence, social security and in ii) administration and secre- tarial occupations belong to different regimes, sublin- ear/linear and superlinear ones respectively. Other em- ployment categories that are predicted to belong to the superlinear regime, such as education, and others that are difficult to justify either way, such as personal service occupations and sales and customer service occupations, lie in the linear regime. The issue is not unique to the employment data as we would also expect to find different exponents for the area of domestic and non-domestic buildings. The former is part of the basic infrastructure of a city, so its sublinear exponent complies with the predicted regime. The latter on the other hand reflects the economic activity of a city, and could be argued to belong to the superlinear regime. However it can be seen from the table that this is not the case. This might be the result of the rent increasing much faster than the economic outcome, constraining the increase of land use.

In all these examples, arguments for and against the hypothesised regimes can be found, but the very fact that they are contestable, makes assignment of an expected regime very difficult for some indicators.

Table S2 contains the details of the variables in Fig. 2. These are urban indicators measured according to the geographical delimitation of British cities given by the European statistical bureau in terms of LUZ. The table gives the values for the scaling exponent, the confidence interval and the precision of the regression.

Variables that are the outcome of social and economic interactions are predicted to have a superlinear exponent. Patents, income and crime incidents show a linear relationship with population size. Often, data on patent registration can be affected by sparse records, however the data used in this study was accumulated for an eleven year period between 2000 and 2011 to ensure the resulting exponent would not be affected by missing values, so it is unlikely that the unexpected exponent for patents is a result of poor data in this case. Data on household income is model-based (see previous section), and assigned to the resident location rather than the workplace location where the income was earned. This may affect results as individuals may choose to live far from the source of their income, and so the wealth generated by a city may not be captured within its physical bounds. We investigate this further by using a different data source for income allocated at the work place from the Official Labour Market Statistics for 2009. The Annual Survey of Hours and Earnings provides weekly gross income at the local authority level, which is aggregated at the LUZ level. We find that the linearity persists, see Fig. S1, indicating that this behaviour may not be due to the specific assignment. The same dataset contains a residential allocation for weekly household income and this also displays linearity implying that the

...
Table S1: Urban indicators for Larger Urban Zones in Great Britain that are difficult to classify in specific regimes

| Urban indicators for LUZ, UK census 2001 | β     | 95% CI   | R² | n. cities |
|-----------------------------------------|-------|----------|----|-----------|
| Employment: public admin and defence, social security | 0.94  | [0.81,1.08] | 0.92 | 21 |
| Employment in admin and secretarial occupations | 1.05  | [1.01,1.09] | 0.99 | 21 |
| Employment in education | 0.96  | [0.87,1.05] | 0.96 | 21 |
| Employment in personal service occupations | 0.97  | [0.94,1.01] | 0.99 | 21 |
| Employment in sales and customer service occ. | 0.99  | [0.95,1.04] | 0.99 | 21 |
| Area of Domestic Buildings (1000m²)* | 0.89  | [0.84,0.95] | 0.99 | 19 |
| Area of Non Domestic Buildings (1000m²)* | 0.88  | [0.81,0.95] | 0.98 | 19 |
| Total number of cars and vans | 0.94  | [0.90,0.99] | 0.99 | 21 |

* The two Welsh cities in LUZ are excluded, since data for Wales on infrastructure is not available

Figure S1: Scatter plot for income data from the Official Labour Market Statistics for 2009, at the household (left) and workplace (right) level. London is clearly an outlier.

result for LUZ is not a dataset artefact. Scrutinising the possible source of conflict, substantial transfers from big cities to small ones might be the cause, i.e. government employment, unemployment benefit, and pensions. These transfers may wash out the effect of increasing productivity by redistributing wealth, leading to a linear relationship of income with the population size.

A map featuring the cities in E&W as prescribed according to LUZ and their size distribution can be found in Fig. S2.

Maps for new definitions of cities
The maps in Fig. S3 correspond to four realisations of systems of cities for Great Britain, given at four different values of the population density threshold. We observe that at very low density, e.g. ρ = 2prs/ha, cities merge into a common cluster, while at very high density, e.g. ρ = 40prs/ha, only the cores of cities can be identified. The maps in Fig. S4 on the other hand, represent the extension of the city boundaries to include wards from which people commute to work to the predefined clusters. This is a set of realisations for the specific population density cutoff of ρ = 14prs/ha, and a minimum population size of N = 50 × 10^3 individuals. The flow indicates the threshold at which the ward will be included. At least the percentage of the population indicated by the flow has to commute from the ward to the cluster in order to be included into the cluster. Noticeable changes in the configuration of the clusters are observed below the threshold of 50%, indicating that it is seldom the case to encounter wards from which more than the majority commutes to a single cluster. As a result the realisation for a flow of 75% is almost identical to system of cities pre-commuting clustering.

Heatmaps for sensitivity analysis
Heatmaps for variables that have a linear relationship with population size will not exhibit significant changes over variations on density and commuting thresholds. On the other hand, power-law distributions are sensitive to the number of events considered. This is clearly observed in the difference obtained for the heatmaps for different minimum population size cutoffs. For the extreme cases, i.e. a high density and population size cutoff, only a small number of cities are taken into account.
Table S2: Urban indicators in Fig. 2 for LUZ in E&W

| Urban indicators for LUZ, UK census 2001 | $\beta$   | 95% CI       | $R^2$ | n. cities |
|-----------------------------------------|----------|--------------|-------|-----------|
| Employed in agriculture, hunting and forestry | 0.59     | [0.35,0.83]  | 0.58  | 21        |
| Distance to work (km)                   | 0.70     | [0.51,0.88]  | 0.76  | 21        |
| Area of Road (1000m$^2$)*               | 0.82     | [0.71,0.93]  | 0.93  | 19        |
| Area of Domestic Gardens (1000m$^2$)*   | 0.84     | [0.75,0.93]  | 0.96  | 19        |
| Employed in manufacturing               | 0.90     | [0.79,1.02]  | 0.93  | 21        |
| n. Bus stops                            | 0.91     | [0.82,1.01]  | 0.96  | 21        |
| Area of Path (1000m$^2$)*               | 0.94     | [0.84,1.05]  | 0.96  | 19        |
| Employed: process; plant and machine op. | 0.95     | [0.84,1.05]  | 0.95  | 21        |
| Employed in elementary occupations      | 0.95     | [0.91,0.99]  | 0.99  | 21        |
| Area of Rail (1000m$^2$)*               | 0.95     | [0.76,1.15]  | 0.86  | 19        |
| Employment in construction              | 0.97     | [0.92,1.02]  | 0.99  | 21        |
| All people aged 16-74 in employment     | 0.99     | [0.96,1.02]  | 1.00  | 21        |
| All household spaces                    | 1.00     | [0.99,1.01]  | 1.00  | 21        |
| Consumption of domestic electricity     | 1.06     | [0.95,1.16]  | 0.96  | 21        |
| Employed in financial intermediation    | 1.25     | [1.13,1.36]  | 0.96  | 21        |
| n. Train Stations                       | 1.12     | [0.86,1.38]  | 0.81  | 21        |
| Employed in real estate, business activities | 1.06   | [0.93,1.20]  | 0.94  | 21        |
| Total income (weekly)                   | 1.03     | [0.96,1.10]  | 0.98  | 21        |
| Employment in associate prof and technical occ. | 1.01   | [0.94,1.08]  | 0.98  | 21        |
| Employed as managers and senior officials | 1.01   | [0.92,1.09]  | 0.97  | 21        |
| Employed in professional occupations    | 1.01     | [0.89,1.12]  | 0.94  | 21        |
| Employed in hotels and restaurants      | 0.97     | [0.94,1.01]  | 0.99  | 21        |
| Crime incidents                         | 0.95     | [0.77,1.13]  | 0.70  | 21        |
| Total number of Patents (2000-2011)     | 0.95     | [0.64,1.26]  | 0.68  | 21        |
| Employment in skilled trades occupations | 0.92     | [0.86,0.98]  | 0.98  | 21        |

* The two Welsh cities in LUZ are excluded, since data for Wales on infrastructure is not available
The plot in Fig. S5 shows how the number of cities considered varies for the different density thresholds and population size cutoffs. The heatmap in Fig. S6 corresponds to distance to work, and it is an example of economies of scale. Its scaling exponent is clearly sensitive to city limits and population size cutoffs. Note the transition for the value of the exponent for the heatmap for a minimum population size of $10^5$, between density thresholds 14 and 15 persons per hectare. And in the heatmap for $N \geq 150 \times 10^3$, a transition to the superlinear regime is observed for very high densities. Fig. S5 shows how these extreme cases are based on only a very few cities.

For all the variables tested, only employment in agriculture, hunting and forestry, and distance to work presented economies of scale. Many other variables that should belong to this category, showed a behaviour more in accordance with linearity: area of roads, paths and domestic buildings; and employment in elementary occupations, manufacturing and construction, see Fig. S7.

Increasing returns effects are also non-existent or very weak for most of the variables. Fig. S8 shows employment categories whose exponent should belong to the superlinear regime, since they require high specialisation, and reflect economic activity. It is expected that the bigger the city, the higher the rate of people employed in skilled tasks. This means that one should find a more than proportional dependency of people employed in these categories with city size. Hence the scaling exponent should lie in the superlinear regime. Nevertheless, Fig. S8 indicates that this is not the case. Such a behaviour occurs only for $\rho \gg 14\text{prs/ha}$, regardless of the commuting flow.

Sensitivity to distribution of cities

There is no consensus on the size a settlement needs to be in order to be considered a city, varying from at least $10^4$ individuals, to agglomerations of at least $10^5$ peo-
Figure S4: Realisations of cities at fixed density cutoff $\rho = 14$prs/ha and minimum population size of $50 \times 10^3$ individuals for a selection of several commuting flow thresholds

Figure S5: This plot shows the number of clusters considered as cities after applying a minimum population size constraint.

Figure S6: Heatmap for distance to work for different minimum population size thresholds: no cutoff, $10^4$, $10^5$ and $150 \times 10^3$.

[9] D. Pumain, F. Paulus, C. Vecchiani-Marcuzzo, and J. Lobo. An evolutionary theory for interpreting urban scaling laws. Cybergeo, 343, 2006.

References
[1] http://www.neighbourhood.statistics.gov.uk/dissemination/
[2] http://www.ons.gov.uk/
[3] http://www.gro-scotland.gov.uk/statistics/.
[4] http://www.ipo.gov.uk
[5] http://research.dse.gov.uk/asd/frs/2001_02/index.php?page=intro
[6] http://research.dse.gov.uk/asd/frs/2001_02/pdonly/frs_2001_02_report.pdf
[7] http://www.homeoffice.gov.uk
[8] http://esa.un.org/unup/index.html.
Figure S7: Each row presents 4 heatmaps for variables that should display economies of scale, for different minimum population size thresholds: no cutoff, $10^4$, $10^5$ and $150 \times 10^3$. A) Infrastructure variables: area of roads, paths and domestic buildings; B) employment categories: elementary occupations, manufacturing and construction.
Figure S8: Heatmap for observables that should belong to the superlinear regime. The exponent however becomes significantly superlinear only when $\rho \gg 14$. 
Figure S9: A) Each point in the figure gives the mean value and standard deviation of $\beta$ for a specific variable. Two distributions are given: blue, when no minimum cutoff is applied; and red, for a minimum cutoff of $150 \times 10^3$ individuals. There is a clear mismatch between the two distributions, otherwise blue and red points would overlap. B) This plot shows the disparity of the scaling exponent at different population size cutoffs. Different values for the mean of $\beta$ over the whole parameter space are recorded. Error bars correspond to the standard deviation. C) Each plot highlights the variability of the scaling exponent, its value and confidence interval, with respect to the different population size cutoffs.