Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 363 (2006) 96–103
PHYSICA A
www.elsevier.com/locate/physa

Scaling laws in urban supply networks

Christian Kühnert[a,*], Dirk Helbing[a,b], Geoffrey B. West[c]

[a]Institute for Transport & Economics, Dresden University of Technology, Andreas-Schubert-Str. 23, 01062 Dresden, Germany
[b]Collegium Budapest—Institute for Advanced Study, Szentháromság u. 2, 1014 Budapest, Hungary
[c]The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Available online 17 February 2006

Abstract

In previous work, it has been proposed that urban structures may be understood as a result of self-organization principles. In particular, researchers have identified fractal structures of public transportation networks and land use patterns. Here, we will study spatial distribution systems for energy, fuel, medical, and food supply. It is found that these systems show power-law scaling as well, when the number of "supply stations" is plotted over the population size. Surprisingly, only some supply systems display a linear scaling with population size. Others show sublinear or superlinear scaling. We suggest an interpretation regarding the kind of scaling law that is expected in dependence of the function and constraints of the respective supply system.

Keywords: Supply network; Scaling; Power law; Urban agglomerations

1. Introduction

The classical view of the spatio-temporal evolution of cities in developed countries is that urban spaces are the result of (centralized) urban planning. After the advent of complex systems' theory, however, people have started to interpret city structures as a result of self-organization processes. In fact, although the dynamics of urban agglomerations is a consequence of many human decisions, these are often guided by optimization goals, requirements, or boundary conditions (such as topographic ones). Therefore, it appears promising to view urban planning decisions as results of the existing structures and upcoming ones (e.g. when a new freeway will lead closeby in the near future). Within such an approach, it would not be surprising anymore if urban evolution could be understood as a result of self-organization.

In fact, already in the 19th century, it was proposed that the migration streams between cities can be understood by a "gravity law" [1,2], according to which the relevant variables for migration activities are the population sizes of the origin and destination town and the distance between them. According to this, city growth could be solely understood as a result of the birth and death rates and migration activities [3].

Maybe even more intriguing is the existence of Zipf's law [4], according to which the population sizes of cities are inversely proportional to their rank. It also implies that the number of cities of a given size is

*Corresponding author.
E-mail address: kuehnert@vwi.tu-dresden.de (C. Kühnert).

0378-4371/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2006.01.058

approximately inversely proportional to their size N, implying a power law

$$p(N) \propto N^{-1} \tag{1}$$

of the size distribution. Among the many different approaches trying to explain Zipf's law (e.g. Refs. [5–7]), the one by Gabaix [8] surprises by its simplicity. According to him, the simplest stochastic model with multiplicative noise $\xi_i(t)$, namely

$$\frac{dN_i}{dt} = [A + \xi_i(t)]N_i(t) \tag{2}$$

is able to generate Zipf's distribution. In agreement with "Gibrat's law" [9,10], it assumes that the growth rates $A_i(t) = A + \xi_i(t)$ are stochastically distributed and varying around a characteristic value A independent of the (population) size $N_i(t)$ of a city i. These days, scientists attempt to explain the different attractiveness of cities, i.e., the time-dependent values $A_i(t)$ through other variables such as productivity, natural resources, quality of life, etc.

Another interesting point is the discovery of self-similar structures in urban systems. For example, Christaller suggested a theory of central places [11], according to which cities are (self-)organized in a hierarchical way: metropoles are surrounded by a hexagonal "ring" of medium-sized cities, each of these by a ring of smaller towns, etc. Later on, the Collaborative Research Project "Natural Constructions"[1] studied urban patterns from the perspective of self-organization theory. For example, street systems were compared with the water transport systems within leaves of plants and analyzed for quantities such as the reachability and average detours. Moreover, public transportation systems were found to show self-similar, fractal features [12]. The same applies to urban boundaries and urban sprawl [7,12–15]. Usage patterns of public transport networks [16] and road networks [17] show power-law distributions as well.

The EU project ISCOM involving biological physicists, geographers, and scientists of other disciplines continues this line of research and goes beyond previous results by trying to identify the mechanisms behind self-similar urban structures and their scaling laws. This is encouraged by the successful derivation of empirically observed scaling laws for quantities like metabolic rates, maximal population growth, life-spans, gross photosynthetic rates, trunk diameters, etc. in biological systems [18–20]. The underlying theory is based on the determination of self-similar, tree-like arterial and other supply systems, which minimize energy dissipation under the constraint of space-covering supply.

In this contribution, we will focus on the question whether scaling laws exist for urban supply systems. If yes, what are their exponents? In Section 2, we start with describing our data sources. Section 3 discusses our data evaluation procedure and compares different statistical evaluation procedures such as linear and logarithmic binning. Section 4 presents the scaling laws found for different supply systems. Finally, Section 5 summarizes our results and tries to give an interpretation of the findings.

2. Data sources

We have studied supply systems of cities in different European countries and analyzed, on the one hand, variables related to the electric energy supply in Germany and, on the other hand, data about so-called "Points of interest" of several European countries.

Our investigation of energy data is based on information by the German Electricity Association [21] for the time period of January to December 2002. The electric energy relates to the power plant of the respective commune, while the population size is taken for the administrative area of the city or municipality the power plant is located in. The population sizes N_i for the population spatial units stem from the German Federal Statistical Office [22] and are based on census data with cut-off date December 2001. Within the countries of investigation, the number of inhabitants of cities usually does not vary dramatically, so that these data should also be applicable to a few years before and after. For large cities such as Berlin, Munich, Bremen, Hannover, or Bremerhaven, there was sometimes more than one enterprise listed. It could also happen that an

[1]It was the "Sonderforschungsbereich 230" of the German Research Foundation (DFG), in which urban planners and architects, construction engineers, biologists, physicists, and philosophers were involved.

superregional supplier had its headquarter in a small city. Such cases have been removed from the data set to avoid outliers.

Moreover, data of urban supply networks have become increasingly available during the last years. In this paper, we have evaluated point-of-interest (POI) data of the year 2002 provided by Teleatlas©. These data were gathered for commercial route guidance and geo-information systems. A potential problem for our evaluation, however, was the different definition of the administrative areas in different countries. In some countries, they were almost on the level of municipalities, in other countries they contained larger areas. To have comparable data, we included in our evaluation only data of statistical areas with more than 50,000 inhabitants and population densities greater than 1000 people per square kilometer. The first condition excluded the small towns and villages as they often represent administrative units rather than naturally grown cities. The second restriction excluded large rural areas.

Another potential problem was that, while some POI variables Y_i had a coverage of over 90% according to information by Teleatlas©, the coverage for others was not complete. We have, therefore, focussed on data sets which were large enough and either close to complete or a good statistical representation. In contrast, we have excluded variables and countries, for which the data appeared to be a non-representative sample or for which the sample was too small (as for countries like Luxembourg or Belgium). Both was reflected by a large data scattering, which discouraged to make any fit to the data, and by large confidence intervals that did not allow for any reasonable conclusions. Generally speaking, we neglected variables if the square of the correlation coefficient regarding the population of the cities was smaller than 0.75.

3. Data evaluation procedure

In our data evaluation, we started with an ordinary least-squares method to a double-logarithmic representation of the data. The data used to scatter in a cone-like manner (see Fig. 1a), which raised questions whether other fit functions could be used as well. However, when we applied a logarithmic binning method [23], which appears to be more appropriate for empirical power-law distributions, the double-logarithmic representation of the number Y_i of supply units over population size N_i (i.e., the plot of $\log(Y_i)$ over $\log(N_i)$) showed that straight lines were usually an excellent fit (see Fig. 1b). This is clearly in favor of power laws

$$Y_i = Y_0 N_i^{-\alpha}, \tag{3}$$

with a scaling exponent α and some constant prefactor Y_0. Both, Y_0 and α depend on the respective supply system under consideration.

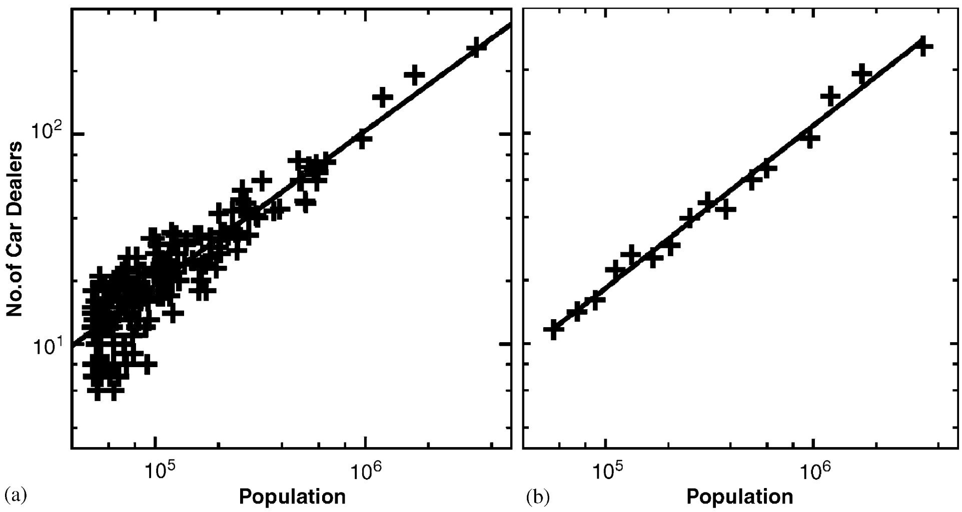

Fig. 1. Double-logarithmic representation of the number of car dealers as a function of the population size of cities in Germany, when displaying (a) the original data and (b) logarithmically binned data. The solid lines correspond to the respective linear regressions.

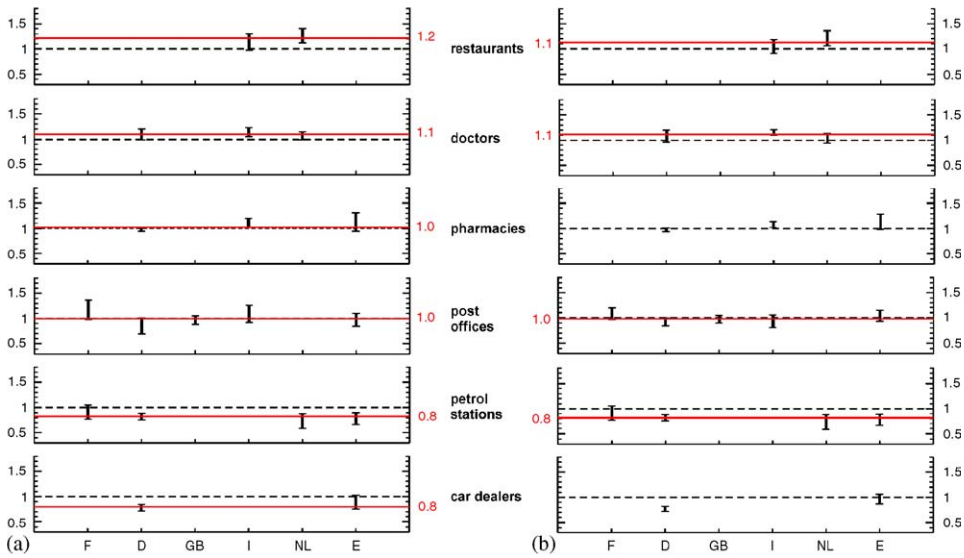

Fig. 2. Scaling exponents and confidence intervals for different supply systems and countries (a) when double-logarithmically scaled original data are used, (b) when the data are logarithmically binned before the regression is made. The countries are: France (F), Germany (D), Great Britain (GB), Italy (I), The Netherlands (NL), and Spain (E).

In the following figures, we will use the logarithmic binning method, which gives a more intuitive representation of the respective underlying relationship. In some sense, it first averages over data points and then fits to averaged data. This is actually fully in the tradition of classical binning methods and compensates for the fact that one would have some overcrowded bins and many empty ones.

Nevertheless, it is not quite clear whether the ordinary linear regression method or the semi-logarithmic binning approach would deliver more reliable estimates of the scaling exponents and their confidence intervals. However, our conclusions for both methods were consistent (see Fig. 2): the average scaling exponents for different countries were about the same, when the same variables were compared. While the logarithmic binning method tended to produce somewhat smaller confidence intervals, ordinary double-logarithmic regression produced intersecting confidence intervals for different countries. The intersection area of the confidence intervals actually allows for a quite precise determination of a scaling exponent, if the same (a "universal") exponent is assumed for all countries (see Fig. 2).

Sometimes there were a lot of bins that were empty or contained the data points of a small number of cities only. Nevertheless, all plots were produced with 20 bins. When we used less bins, the fraction of sparsely filled bins remained almost the same, but the confidence intervals increased.

When we applied the statistical procedures of the software package stata©, we checked the precondition of homoskedasticity for the ordinary least-square regression (OLS) by a residual-versus-fit plot and by a Cook–Weisberg/Breusch–Pagan test. If this assumption was not fulfilled, we applied the feasible general least-square regression (FGLS), which uses a correction for heteroskedasticity according to Long and Ervin.

4. Scaling laws of urban supply systems

In the following, we present the evaluation results of our supply system data sets. Due to their greater clarity, we have plotted the data using the logarithmic binning method. The plots for electrical energy supply in Germany are shown in Fig. 3. A summary of the scaling exponents and the 95% confidence intervals (both determined for the original data) is presented in Table 1.

Figs. 4–6 show the plots for the number of petrol stations, post offices, and restaurants. Note that all the plots support a linear relationship between the logarithm of the number of respective "supply stations" and the logarithm of population size. This implies that the dependence of the number of "supply stations" as a function of the population size follows a power law.

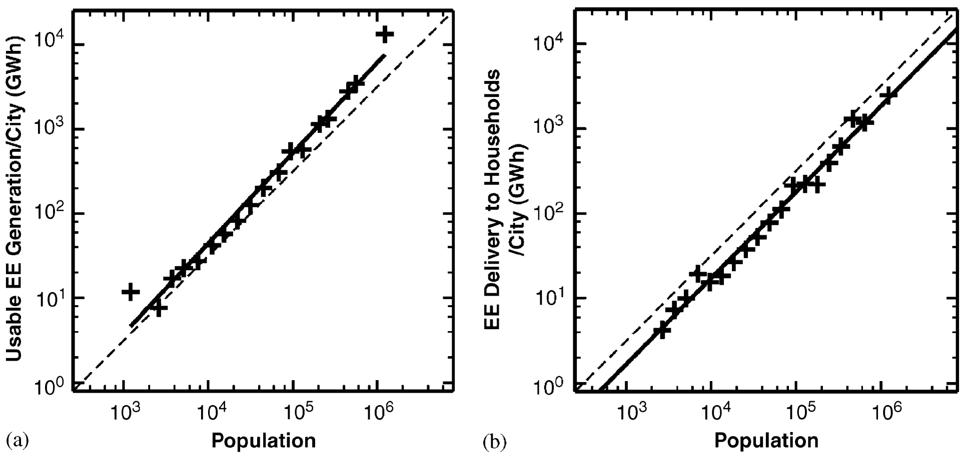

Fig. 3. Double-logarithmic representation of (a) the usable electric energy (EE) and (b) the energy delivery to households, both as a function of the population size of various German cities, after a logarithmic binning method has been applied. The solid lines correspond to the respective linear regression and the dashed lines indicate the slope 1.

Table 1
Scaling exponents and their 95% confidence intervals for different variables of electric energy supply in Germany as a function of population size

Variable	Exponent	95% confidence interval
Usable electric energy	1.1	[1.04; 1.13]
Electric energy delivery to households	1.0	[0.96; 1.06]
Length of low-voltage cables	0.9	[0.81; 0.92]

5. Summary and discussion

In this paper, we have analyzed empirical data of urban supply systems. When the number of "supply stations" is fitted as a function of the population size, one can find power-law distributions for quantities as different as electrical energy, petrol stations, car dealers, hospitals, hospital beds, post offices, pharmacies, doctors, and restaurants. For several European countries, the corresponding power-law exponents and their 95% confidence intervals are summarized in Fig. 2. No data were displayed, if the data sample were too small, not representative, or if the presumptions of the respective statistical tests were violated.

Despite the width of the confidence intervals, one can draw several interesting conclusions:

1. The scaling exponents of different countries were consistent, i.e., of the same order. In fact, the 95% confidence intervals tend to have a common subset, which may be used for a more precise determination of the respective scaling exponent, if universality (i.e., country-independence) is assumed. Statistical analysis of variance tests support this picture.
2. A proportionality of the number of "supply stations" to the population size corresponding to a scaling exponent of 1 is only found for some supply systems. This includes hospitals and hospital beds, post offices, and pharmacies.
3. There were also cases of sub- or superlinear relationships. For example, the scaling exponents for the number of car dealers and petrol stations were smaller than 1 (sublinear case), while the scaling exponent for restaurants was larger than 1 (superlinear case).

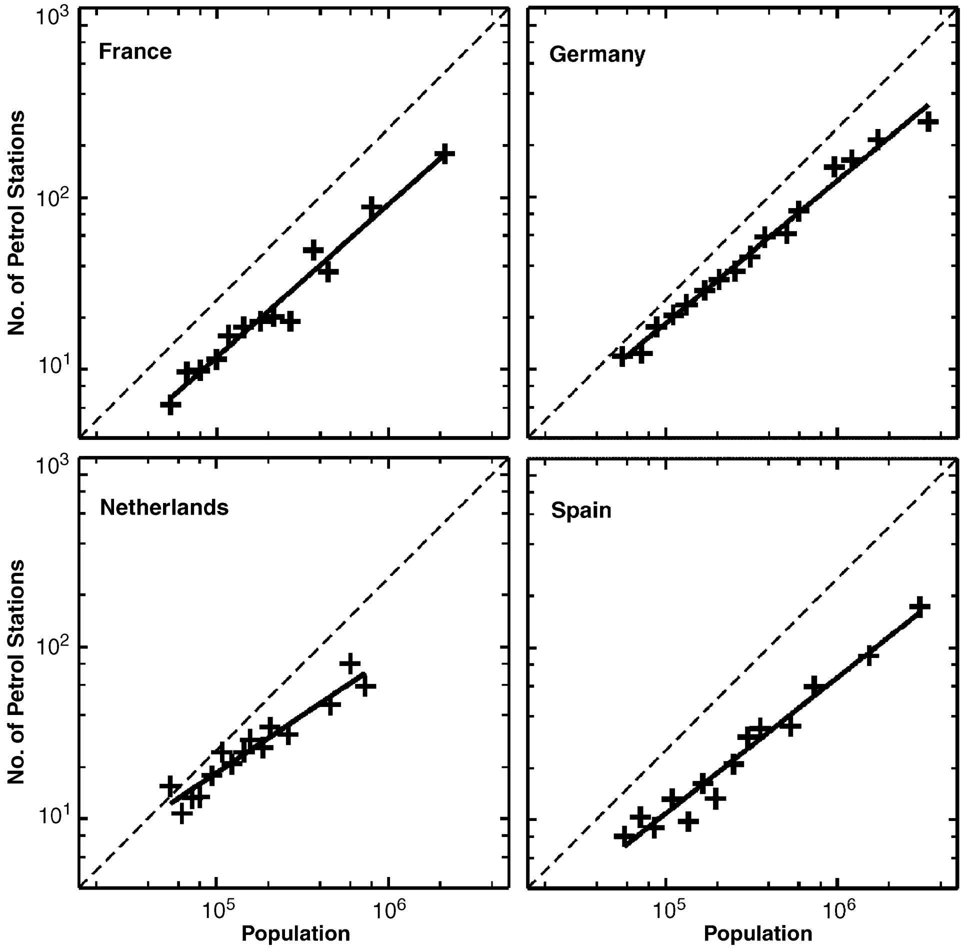

Fig. 4. Double-logarithmic representation of the number of petrol stations as a function of the population size of cities of France, Germany, Netherlands and Spain, after a logarithmic binning method has been applied. The solid lines correspond to the respective linear regression and the dashed lines indicate the slope 1.

What are the reasons for the observed differences in the scaling exponents? The proportionality of post offices, pharmacies and doctors to the population size is probably dictated by comparable individual demands, combined with the requirement of a certain level of reachability (by foot). Moreover, it is often regulated by government. As a consequence, each size class of cities offers approximately the same number of these "supply stations".

Sublinearly scaling quantities such as the number of petrol stations or car dealers indicate an "economy of scales". That is, one such "supply station" serves more people in a larger town and distributes larger quantities (e.g. sells more fuel per month). This is certainly reasonable and typical for material supply systems, which are profitable for large population sizes only or governed by a free market. Sublinear scaling supply systems profit from higher population densities and a more efficient usage of capacities in larger service units (e.g. by better utilization or reduction of the relative statistical variation, etc.). Sublinear scaling is also expected for the number of shopping centers or for polyclinics.

Finally, why do some supply systems scale superlinearly? This concerns, for example, the number of restaurants, but a similar thing seems to be true for museums, theaters, colleges, etc. We identify these as the supply systems satisfying social and communicative needs. That is, information exchange seems to increase more than proportional with the number of inhabitants in a town. The number of patents as a function of the population size [24] and other variables [25] seem to confirm this conclusion.

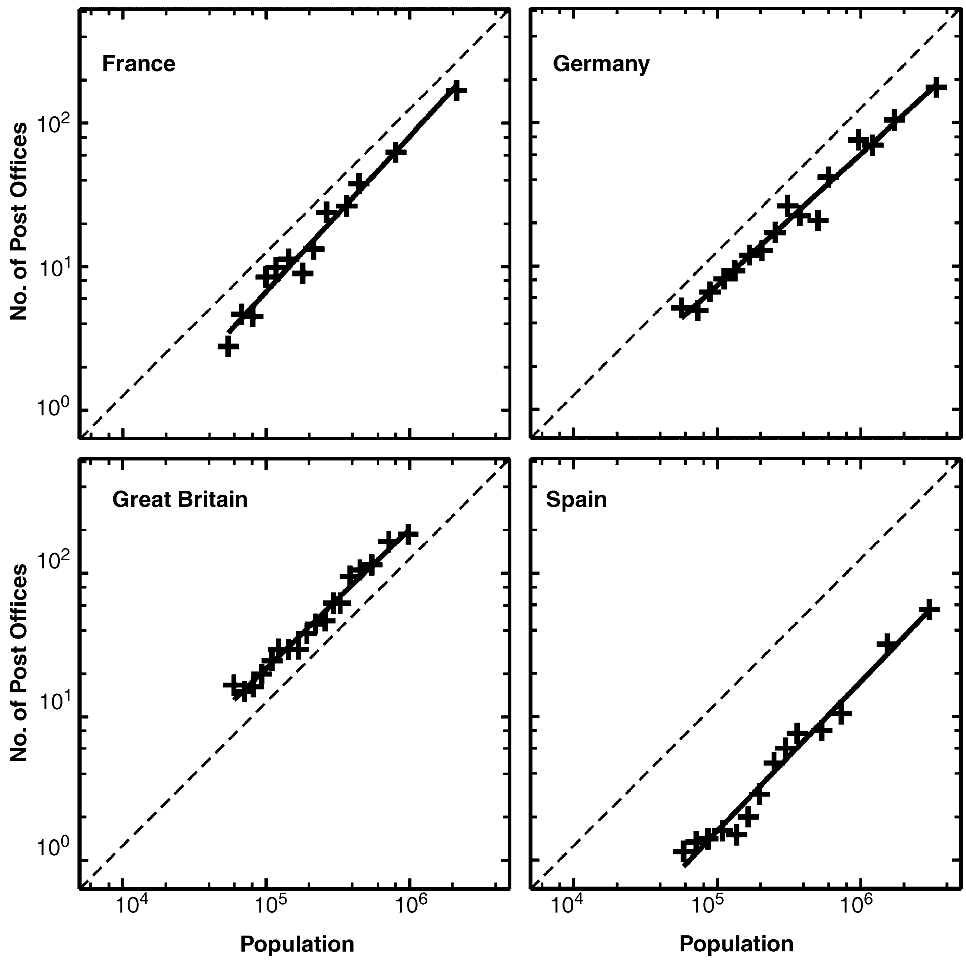

Fig. 5. Double-logarithmic representation of the number of post offices as a function of the population size of cities in France, Germany, Great Britain and Spain, after a logarithmic binning method has been applied. The solid lines correspond to the respective linear regression and the dashed lines indicate the slope 1.

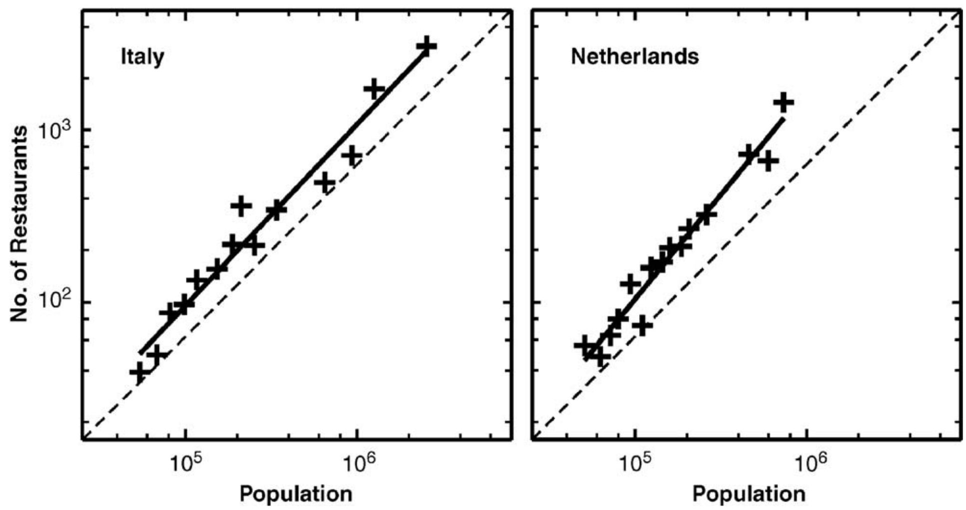

Fig. 6. Double-logarithmic representation of the number of restaurants as a function of the population size of cities of Italy and Netherlands, after a logarithmic binning method has been applied. The solid lines correspond to the respective linear regression and the dashed lines indicate the slope 1.

Acknowledgments

This study has been partially supported by the EU project ISCOM and the DFG project He2789/5-1. The authors would like to thank Luis Bettencourt, Jose Lobo, and Denise Pumain for inspiring discussions. They are also grateful to the VDEW for the data of the German electricity suppliers. Moreover, they acknowledge the data evaluation efforts of Markus Winkelmann in a previous stage of this study.

References

[1] E. Ravenstein, The Geographical Magazine III (1876) 173–177, 201–206, 229–233.
[2] G.K. Zipf, Am. Sociol. Rev. 11 (1946) 677.
[3] W. Weidlich, G. Haag (Eds.), Interregional Migration, Springer, Berlin, 1988.
[4] G.K. Zipf, Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology, Addison-Wesley, Cambridge, MA, 1949.
[5] J. Steindl, Random Processes and the Growth of Firms, Hafner, New York, 1965.
[6] H. Simon, Biometrika XLII (1955) 425.
[7] F. Schweitzer (Ed.), Self-Organization of Complex Structures: From Individual to Collective Dynamics, Gordon and Breach, London, 1997.
[8] X. Gabaix, Q. J. Econ. 114 (3) (1999) 739.
[9] R. Gibrat, Les Inégalités Économiques, Librairie du Recueil Sirey, Paris, 1931.
[10] J. Sutton, J. Econ. Lit. XXXV (1997) 40.
[11] W. Christaller, Die zentralen Orte in Süddeutschland, Wissenschaftliche Buchgesellschaft, Darmstadt, 1980.
[12] P. Frankhauser, La Fractalité des Structrures [sic] Urbaines, Anthropos, Paris, 1994.
[13] E. Schaur, Ungeplante Siedlungen/Non-Planned Settlements, Krämer, Stuttgart, 1991.
[14] H.A. Makse, S. Havlin, H.E. Stanley, Nature 377 (1995) 608.
[15] M. Batty, P. Longley, Fractal Cities: A Geometry of Form and Function, Academic Press, London, San Diego, 1994.
[16] J.A. Hołyst, J. Sienkiewicz, A. Fronczak, P. Fronczak, K. Suchecki, Phys. Stat. Mech. Appl. 351 (2005) 167.
[17] S. Lämmer, B. Gehlsen, D. Helbing, Scaling laws in the spatial structure of urban road networks, in this special issue of Physica A.
[18] G.B. West, J.H. Brown, B.J. Enquist, Science 276 (1997) 122.
[19] G.B. West, W.H. Woodruft, J.H. Brown, Proc. Natl. Acad. Sci. USA 99 (2002) 2473.
[20] B.J. Enquist, J.H. Brown, G.B. West, Nature 395 (1998) 163.
[21] Verband der Elektrizitätswirtschaft, Jahresdaten der Stromversorger 2002, VWEW Energieverlag GmbH, Frankfurt am Main, Heidelberg, 2003.
[22] Statistische Ämter des Bundes und der Länder, Statistik lokal. Daten für die Gemeinden und Kreise Deutschlands (CD-ROM) 2003.
[23] O.B. McManus, A.L. Blatz, K.L. Magleby, Pflügers Arch. Eur. J. Physiol. 410 (1987) 530.
[24] D. Strumsky, J. Lobo, L. Fleming, Santa Fe Institute Working Paper 05-02-004, 2005.
[25] D. Pumain, F. Paulus, C. Vacchiani, J. Lobo, An evolutionary theory for urban scaling laws, Cybergeo, 2005, submitted for publication.