

The growth equation of cities

<https://doi.org/10.1038/s41586-020-2900-x>

Vincent Verbavatz^{1,2} & Marc Barthelemy^{1,3✉}

Received: 13 February 2020

Accepted: 1 September 2020

Published online: 18 November 2020

 Check for updates

The science of cities seeks to understand and explain regularities observed in the world's major urban systems. Modelling the population evolution of cities is at the core of this science and of all urban studies. Quantitatively, the most fundamental problem is to understand the hierarchical organization of city population and the statistical occurrence of megacities. This was first thought to be described by a universal principle known as Zipf's law^{1,2}; however, the validity of this model has been challenged by recent empirical studies^{3,4}. A theoretical model must also be able to explain the relatively frequent rises and falls of cities and civilizations⁵, but despite many attempts^{6–10} these fundamental questions have not yet been satisfactorily answered. Here we introduce a stochastic equation for modelling population growth in cities, constructed from an empirical analysis of recent datasets (for Canada, France, the UK and the USA). This model reveals how rare, but large, interurban migratory shocks dominate city growth. This equation predicts a complex shape for the distribution of city populations and shows that, owing to finite-time effects, Zipf's law does not hold in general, implying a more complex organization of cities. It also predicts the existence of multiple temporal variations in the city hierarchy, in agreement with observations⁵. Our result underlines the importance of rare events in the evolution of complex systems¹¹ and, at a more practical level, in urban planning.

Constructing a science of cities has become a crucial task for our societies, which are growing ever more concentrated in urban systems. Better planning could be achieved with a better understanding of city growth and how it affects society and the environment¹². Various important aspects of cities such as urban sprawl, infrastructure development or transport planning depend on the population evolution over time, and multiple theoretical attempts have been made in order to understand this crucial phenomenon.

Growth of cities and Zipf's law

So far, most research in city growth has been done with the idea that the stationary state for a set of cities is described by Zipf's law. This law is considered to be a cornerstone of urban economics and geography³, and states that the population distribution of urban areas in a given territory (or country) displays a Pareto law with exponent equal to 2 or, equivalently, that the city populations sorted in decreasing order versus their ranks follow a power law with exponent 1. This alleged regularity through time and space is probably the most striking fact in the science of cities and for more than a century has triggered intense debate and many studies^{1,2,5,10,13–28}. This result characterizes the hierarchical organization of cities, and in particular it quantifies the statistical occurrence of large cities. Zipf's law implies that in any country, the city with the largest population is generally twice as large as the next largest, and so on. It is a signature of the very large heterogeneity of city sizes and shows that cities are not governed by optimal considerations that would lead to one unique size but, on the contrary, that city sizes are broadly distributed and follow some sort of hierarchy¹⁶. The empirical

value of the Pareto exponent informs us about the hierarchical degree of a system of cities: a large value of the exponent corresponds to a more equally distributed population among cities, and, vice versa, for small exponent values the corresponding system of cities is very heterogeneous with a few megacities.

Studies in economics have suggested that Zipf's law is the result of economic shocks and random growth processes^{6–8}. Gabaix¹⁰ proved in a seminal paper that Gibrat's law of random growth⁹—which assumes a population growth rate independent of the size of the city—can lead to a Zipf law with exponent 1, at the expense of the additional and untested assumption that cities cannot become too small. This model remains the most accepted paradigm to understand city growth. Since then, it has also been understood using simplified theoretical models (without any empirical arguments) that migrations from other cities or countries are determinant in explaining random growth²⁹. However, although most of these theoretical approaches focus on explaining Zipf's law with exponent 1, recent empirical studies^{3,4}, supported by an increasing number of data sources, have questioned the existence of such a universal power law and have shown that Zipf's exponent can vary around 1 depending on the country, the time period, the definition of cities used or the fitting method^{13,21,30,31} (we illustrate this in Extended Data Fig. 1, showing that no universal result for the population distribution is observed), leading to the idea that there is no reason to think that Zipf's law holds in all cases³².

Beyond understanding the stationary distribution of urban populations lies the problem of their temporal evolution. As already noted⁵, the huge number of studies regarding population distribution contrasts with the few analyses of the time evolution of cities.

¹Institut de Physique Théorique, Université Paris-Saclay, CNRS, CEA, Gif-sur-Yvette, France. ²École des Ponts ParisTech, Champs-sur-Marne, France. ³Centre d'Etude et de Mathématique Sociales, CNRS/EHESS, Paris, France. ✉e-mail: marc.barthelemy@ipht.fr

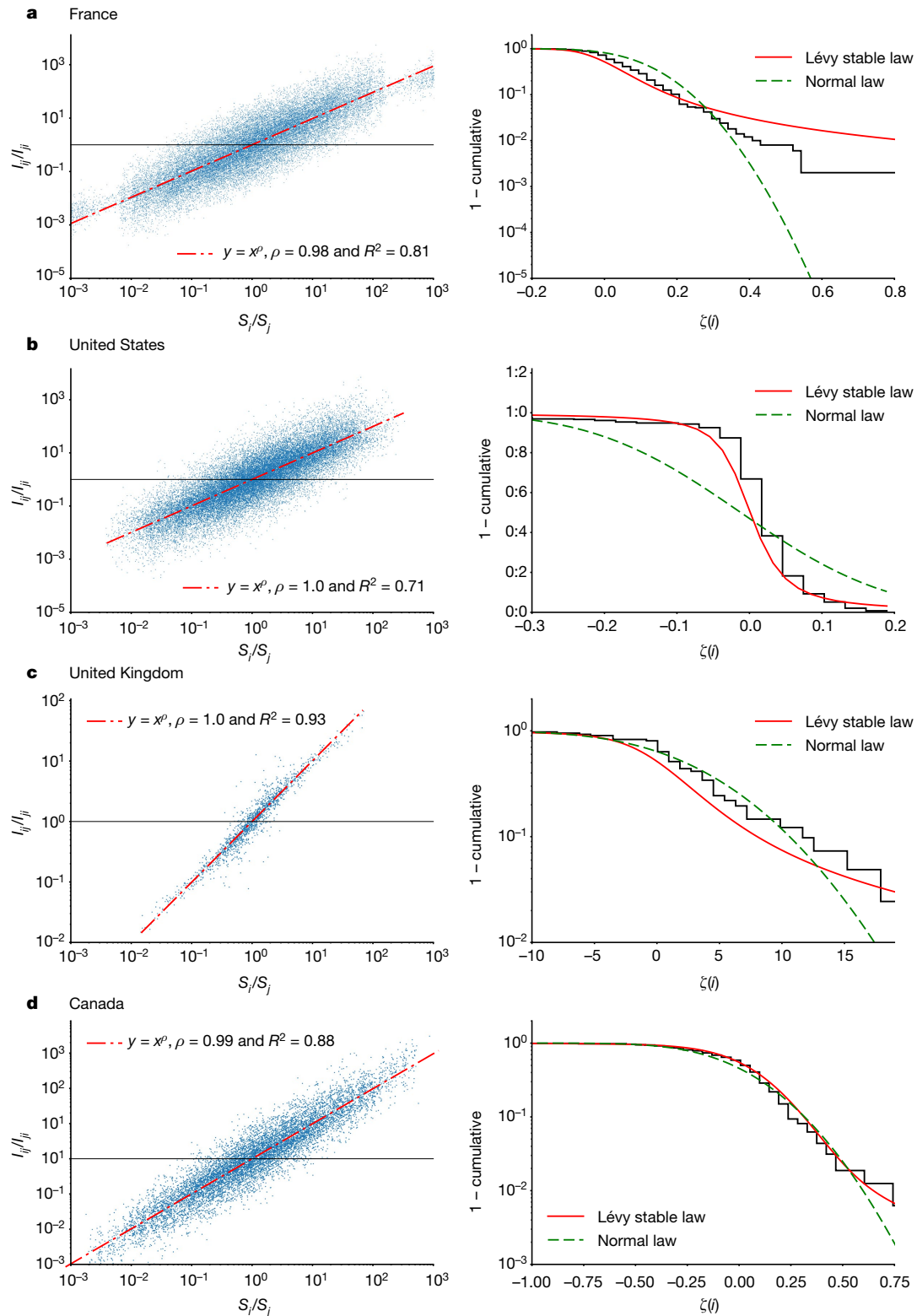


Fig. 1 | Migration flow analysis. **a–d**, Analysis for France (**a**), the USA (**b**), the UK (**c**) and Canada (**d**). Left, migration-rate ratio versus the ratio of populations. The straight line is a power-law fit that gives an exponent equal to one. Right, empirical right-cumulative distribution function of renormalized migrations

flows ζ_i compared to Lévy (continuous red lines) and normal distributions (green dashed lines). See Extended Data Fig. 4 for the left-cumulative distribution function.

As discussed in that same work⁵, cities and civilizations rise and fall many times on a large range of time scales, and Gabaix’s model is both quantitatively and qualitatively unable to explain these specific chaotic dynamics.

Therefore, a model able to simultaneously explain observations about the stationary population distribution and the temporal dynamics of systems of cities is missing. In particular, we are not at this point able to identify the causes of the diversity of empirical

Table 1 | Estimates of parameter α

Dataset	MLE	Kolmogorov–Smirnov test	Log-moments	Hill
France, 2003–2008	1.43 ± 0.07	1.2 < α < 1.8	1.3	1.4 ± 0.3
USA, 2013–2017	1.27 ± 0.07	1.15 < α < 1.20	1.2	Inconclusive
UK, 2012–2016	1.32 ± 0.26	Inconclusive	1.0	1.2 ± 0.8
Canada, 2012–2016	1.69 ± 0.12	Inconclusive	1.9	1.4 ± 0.6

We used four different methods of estimation: maximum-likelihood estimation, Kolmogorov–Smirnov test, log-moments and Hill estimates (see, for example, ref. 42).

observations about the hierarchical organization of cities, the occurrence of megacities, and the empirical instability in city dynamics seen in the births and deaths of large cities on short time scales. In this respect, we do not need just a quantitative improvement of models but a shift of paradigm.

In this paper, we show that city growth is dominated by rare events—namely large interurban migratory shocks—rather than by the average growth rate. Rare but large positive or negative migratory flows can destabilize the hierarchy and the dynamics of a city on very short time scales, leading to the disordered dynamics of cities observed throughout history. On the basis of an empirical analysis of migrations flows in four countries, in the following we derive a stochastic equation of city growth that is able to explain empirical observations of the statistics and temporal dynamics of cities.

Deriving the equation of city growth

To understand city growth, we require a robust, bottom-up approach, starting from elementary mechanisms governing the evolution of cities. Without loss of generality, the growth dynamics of a system (such as a country) of cities i of size S_i can be decomposed into the sum of an interurban migration term between metropolitan areas and an ‘out-of-system’ term that combines other sources of growth: natural growth (births and deaths) and migrations that do not occur within the system of cities (international migrations and exchanges with smaller towns). We denote by $N(i)$ the set of neighbours of city i , that is, those that exchange a non-zero number of inhabitants. Using the four recent datasets of migrations that we use here (USA, 2012–2017; France, 2003–2008; England and Wales (for simplicity, UK), 2012–2016; Canada, 2012–2016) we find for France and the USA that $|N(i)| \propto S_i^\gamma$, where $\gamma \approx 0.5$ (Extended Data Fig. 2). The British and Canadian datasets are fully connected, leading to $\gamma = 0$. The time (t) evolution of the population size S_i can then be written as

$$\frac{\partial S_i}{\partial t} = \eta_i S_i + \sum_{j \in N(i)} J_{j \rightarrow i} - J_{i \rightarrow j}, \quad (1)$$

where the quantity η_i is a random variable accounting for the ‘out-of-system’ growth of city i ; the data show that η_i is Gaussian-distributed (Extended Data Fig. 3). The flow $J_{i \rightarrow j}$ is the number of individuals moving from city i to city j during a period of time dt . If there is an exact balance of migration flows ($J_{i \rightarrow j} = J_{j \rightarrow i}$), the equation becomes equivalent to Gibrat’s model⁹, which predicts a log-normal distribution of populations.

Starting from this general equation (1) is very natural as it amounts to writing the balance of births, deaths and migrations; however—as is often the case when using very general, basic equations—it is difficult to use for making predictions. Simplifications of this equation have been proposed²⁹, wherein various assumptions (such as the gravity model for migration, for example) lead to Gibrat’s model, but miss the very large fluctuations of migrations—as we will see below, this is a crucial ingredient. We also note that this general stochastic equation (1) was discussed in another context³³ and is a central object in the

Table 2 | Estimates of parameters for the four datasets

Dataset	γ	ν	$\beta = \nu + \gamma/\alpha$	β_{measured}
France, 2003–2008	0.55 ± 0.06	0.4 ± 0.3	0.8 ± 0.4	0.75 ± 0.07
USA, 2013–2017	0.41 ± 0.05	0.4 ± 0.4	0.7 ± 0.5	0.93 ± 0.07
UK, 2012–2016	0	0.7 ± 0.3	0.7 ± 0.3	0.51 ± 0.05
Canada, 2012–2016	0	0.5 ± 0.4	0.5 ± 0.4	0.78 ± 0.06

We observe good agreement between the measured and predicted values of β in the USA and France (see Supplementary Information for details about these estimates). The British and Canadian datasets are small and hence fully connected (implying $\gamma = 0$), and are very noisy.

statistical physics of disordered systems. With regard to cities, the migration flow $J_{i \rightarrow j}$ depends a priori (and at least) on the populations S_i and S_j and the distance d_{ij} between cities i and j . Using a standard gravitational model^{34,35}, we show that for France and the USA, the dominant contribution to $J_{i \rightarrow j}$ comes from the populations and that the role of distance appears as a second-order effect (see Supplementary Information for details). This result suggests that the $J_{i \rightarrow j}$ term can be represented by a variable of the form $I_0 S_i^\mu S_j^\nu x_{ij}$, where the random variables x_{ij} have an average equal to 1 and encode the noise as well as multiple other effects, including distance. We denote by $I_{ji} = J_{i \rightarrow j}/S_i$ the probability per unit time and per capita of moving from city i to city j . The left panel of Fig. 1 shows that the ratio I_{ij}/I_{ji} versus the ratio of populations S_i/S_j displays, on average, linear behaviour. This implies that $\mu = \nu$, and that we have, on average, a sort of detailed balance $\langle J_{i \rightarrow j} \rangle = \langle J_{j \rightarrow i} \rangle$ (where the angled brackets here denote the average over cities), but that crucially, fluctuations are non-zero. More precisely, if we denote by $X_{ij} = (J_{j \rightarrow i} - J_{i \rightarrow j})/I_0 S_i^\nu$, we observe that these random variables X_{ij} are heavy-tailed—that is, they are distributed according to a broad law that decreases asymptotically as a power law with exponent $\alpha < 2$ (see Supplementary Information for more details and empirical evidence). The sum in the second term of the right-hand side of equation (1) can then be rewritten as

$$\sum_{j \in N(i)} J_{j \rightarrow i} - J_{i \rightarrow j} = I_0 S_i^\nu \sum_{j \in N(i)} X_{ij}, \quad (2)$$

and, according to the generalized version of the central limit theorem³⁶ (assuming that correlations between the variables X_{ij} are negligible), the random variable

$$\zeta_i = \frac{1}{|N(i)|^{1/\alpha}} \sum_{j \in N(i)} X_{ij}$$

follows a Lévy stable law L_α with parameter α (for large enough $N(i)$). This is empirically confirmed in Fig. 1 (right panel): French, US, British and Canadian data are better fitted by a Lévy stable law than by any other distribution and the estimates of α (using different methods) are given in Table 1. We are led to the conclusion that the growth of systems of cities is governed by a stochastic differential equation with two independent noises, which reads as follows

$$\frac{\partial S_i}{\partial t} = \eta_i S_i + D S_i^\beta \zeta_i, \quad (3)$$

where $D \propto I_0$, $\beta = \nu + \gamma/\alpha$ and η_i is a Gaussian noise with mean the average growth rate r and a dispersion σ . This is the growth equation of cities that governs the dynamics of large urban populations; it is our main result here. In equation (3) both noises are uncorrelated and multiplicative, and Itô’s convention here seems to be more appropriate than Stratonovich’s³⁷ because population sizes at time t are computed independently from interurban migration terms at time $t + dt$. Estimates for the various parameters together with the prediction for the value of β are given in Table 2.

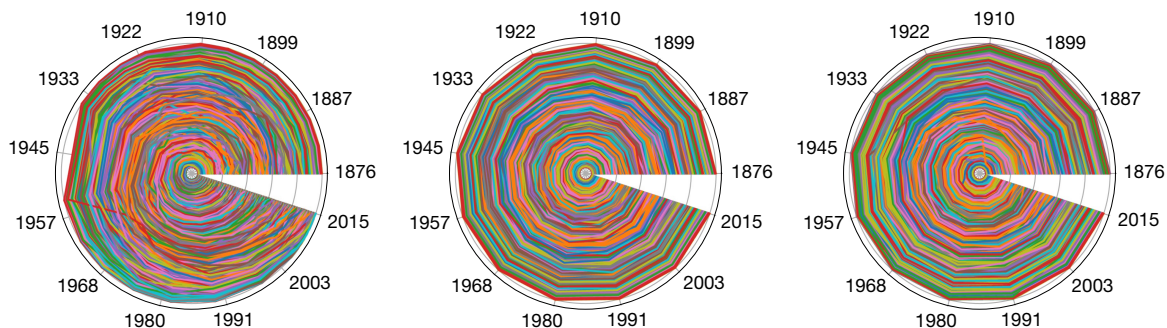


Fig. 2 | Rank clocks for France. We compare the real dynamics of the 500 largest French cities between 1876 and 2015 (left) to Gabaix’s statistical prediction (middle) and to our statistical prediction (right). On the clocks, each

line represents a city rank over time where the radius is given by the rank and the angle by time. In this representation, the largest city is at the centre and the smallest at the edge of the disk.

The central limit theorem, together with the broadness of interurban migration flow, enables us to show that many details in equation (1) are unnecessary and that the dynamics can be described by the more universal equation (3). We conclude that starting from equation (1) is thus less useful than previously thought. The importance of migrations has been previously noted²⁹, but in that work the authors derived a stochastic differential equation with multiplicative Gaussian noise, which we show here to be incorrect: we indeed have a first term with multiplicative noise but also, crucially, we obtain another term that is a multiplicative Lévy noise with zero average. This is a major theoretical shift that is not included in previous studies on urban growth and which has many crucial implications in understanding both the stationary and dynamic properties of cities.

No stationary distribution for cities

Equation (3) governs the evolution of urban populations and analysing it at large times gives indications about the stationary distribution of cities. To discuss the analytical properties of equation (3), we assume that Gaussian fluctuations are negligible compared to the Lévy noise and write $\eta_i \approx r$ (see Extended Data Fig. 5). The corresponding Fokker–Planck equation (with Itô’s convention) can be solved using the formalism of fractional-order derivatives and Fox functions^{38–41}, leading to the general distribution at time t that can be expanded in powers of S as (see Supplementary Information for derivation and complete expressions of all terms):

$$P(S, t) = \sum_{k=1}^{\infty} C_k \frac{a(t)^{-\alpha\beta - \alpha(1-\beta)k}}{S^{1+\alpha\beta + \alpha(1-\beta)k}} \quad (4)$$

where C_k is a prefactor that is a function of α , β and k and independent of t and S , and where $a(t) \propto \left[\frac{r/D^\alpha}{(e^{r\alpha(1-\beta)t} - 1)} \right]^{-1/\alpha(1-\beta)}$ decreases exponentially at large times. This expansion shows that the probability distribution of city sizes is dominated at large S by the order $k = 1$ and converges towards a Pareto distribution with exponent $\alpha \neq 1$. The speed of convergence towards this power law can be estimated with the ratio $\lambda(S, t)$ of the first and second terms of the expansion equation (4) and leads to:

$$\lambda(S, t) = \frac{D^\alpha}{r} \left(\frac{\bar{S}(t)}{S} \right)^{\alpha(1-\beta)} \quad (5)$$

where $\bar{S}(t)$ is the mean city size. If $\lambda(S) \geq 1$, the α -exponent regime is not valid in the right tail with threshold S at time t . Estimates of α and β for the four datasets show that finite-time effects are very important in all cases and that a power-law regime is only reached for unrealistically large city sizes (see discussion in Supplementary Information). Hence, the range of city sizes for which we can observe a power-law distribution may not exist in practice and there is no reason in general to observe

Zipf’s law or any other stationary distribution. We also note that from equation (4) there is a scaling of the form $P(S, t) = \frac{1}{S} F\left(\frac{S}{\bar{S}(t)}\right)$ with a scaling function F that depends on the country. We confirmed this scaling form for France (the only country for which we had sufficient data); details can be found in Supplementary Information (see also Extended Data Fig. 6).

In addition, if we perform a power-law fit of the expansion (equation (4)), the upper tail of the city-size distributions may be mistaken for a Pareto tail with a spurious exponent that changes with the definition of the upper tail (Extended Data Fig. 7). This might explain the discrepancies observed in the literature on Zipf’s law. As city sizes increase, the apparent exponent changes and can dramatically deviate from 1, as we initially observe in Extended Data Fig. 1. Following our analysis, the apparent exponent should converge towards the value given by α , as is indeed observed in, for example, France ($\alpha = 1.4$) and the USA ($\alpha = 1.3$).

Dynamics: splendour and decline of cities

The validity of our model (equation (3)) can be further tested on the dynamics of systems of cities over large periods of time. This can be done by following the populations and ranks of the system’s cities at different times with the help of ‘rank clocks’, as previously proposed⁵. In that work, it was proven that the micro-dynamics of cities is very turbulent, with many rises and falls of entire cities that cannot result from Gabaix’s model (which is, in essence, Gibrat’s model with a non-zero minimum for city sizes). We show in Fig. 2 the empirical rank clock for France (from 1876 to 2015) and for the results obtained with Gabaix’s model and ours (for the other countries, see Extended Data Fig. 8).

We see that in Gabaix’s model (middle), the city rank is stable on average, and not turbulent: the rank trajectories are concentric and the rank of a city oscillates around its average position. In the real dynamics (left), cities can emerge or die. Very fast changes in rank order can occur, leading to much more turbulent behaviour. In our model (right), the large fluctuations of Lévy noise are able to statistically reproduce such ebbs and flows of cities. More quantitatively, we first compare the average shift per time $d = (\sum_t \sum_{i=1}^N |r_i(t) - r_i(t-1)|) / NT$ over T years and for N cities in the three cases (Table 3) and look at the statistical fluctuations of the rank (see Extended Data Fig. 9): we note that Lévy fluctuations are much more able to reproduce the turbulent properties of the dynamics of cities through time. Indeed, the fast births and deaths of cities—due, for example, to wars, discoveries of new resources, incentive settlement policies, and so on—are statistically explained by broadly distributed migrations and are incompatible with a Gaussian noise. Second, we can compare with the empirical data the predictions of the different models for the time needed to make the largest rank jump (see Extended Data Fig. 10 for France, which typically predicts a duration of order 80 years to make a very large jump). We

Table 3 | Average rank shift per unit time, d

Average shift per unit time, d	Real	Lévy	Gabaix
France, 1876–2015	6.0	6.1	8.0
USA, 1790–1990	4.7	16	27
UK, 1861–1991	4.8	16	25

Parameters for the Lévy model and the Gaussian noise are fitted on the France 2003–2008, USA 2013–2017 and UK 2012–2016 datasets, respectively. The most complete dataset is the French one, with total population for all cities at all times. In the US and UK datasets, only the populations of the largest cities are recorded (top 100 in the USA and top 40 in the UK). This can explain the large discrepancies observed, considering the distance d in both countries. Extended Data Fig. 10 investigates the distribution and fluctuations of the rank over time.

confirm that Gabaix’s model is unable to reproduce these very large fluctuations and that our equation agrees very well with the data.

A new paradigm

In this Article, we build a stochastic equation of growth for cities on the basis of microlevel considerations that is empirically sound and that challenges the paradigm of Zipf’s law and current models of urban growth. We show that microscopic details are irrelevant and that the growth equation obtained is universal. A crucial point in this reasoning is that, although we have on average a sort of detailed balance that would lead to a Gaussian multiplicative-growth process, it is the existence of non-universal and broadly distributed fluctuations of the microscopic migration flows between cities that govern the statistics of city populations. We introduce here a stochastic equation that describes city growth that includes two sources of noise and that predicts an asymptotic power-law regime. However, this stationary regime is not generally reached and finite-time effects cannot be discarded. Our model is also able to statistically reproduce the turbulent micro-dynamics of cities that rapidly rise and fall, in contrast with previous Gaussian-based models of growth⁵.

In addition, our fundamental result exhibits an interesting connection between the behaviour of complex systems and non-equilibrium statistical physics for which microscopic currents and the violation of detailed balance seem to be the rule rather than the exception¹¹. At a practical level, this result also highlights the critical effect of not only interurban migration flows (an ingredient that is not generally considered in urban-planning theories), but also, more importantly, their large fluctuations—which are ultimately connected to the capacity of a city to attract a large number of new citizens. Our approach, which relies in essence on the population budget description and empirical results, provides a solid ground for future research on the temporal evolution of cities, a central problem in urban science.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2900-x>.

1. Zipf, G. K. *Human Behavior and the Principle of Least Effort* (Addison-Wesley, 1949).

- Auerbach, F. Das Gesetz der Bevölkerungskonzentration. *Petermanns Geogr. Mitt.* **59**, 74–76 (1913).
- Arshad, S., Hu, S. & Ashraf, B. N. Zipf’s law and city size distribution: a survey of the literature and future research agenda. *Physica A* **492**, 75–92 (2018).
- Gan, L., Li, D. & Song, S. Is the Zipf law spurious in explaining city-size distributions? *Econ. Lett.* **92**, 256–262 (2006).
- Batty, M. Rank clocks. *Nature* **444**, 592–596 (2006).
- Duranton, G. & Puga, D. in *Handbook of Economic Growth* Vol. 2 (eds Aghion, P. & Durlauf, S.) 781–853 (Elsevier, 2014).
- Córdoba, J.-C. On the distribution of city sizes. *J. Urban Econ.* **63**, 177–197 (2008).
- Rossi-Hansberg, E. & Wright, M. Urban structure and growth. *Rev. Econ. Stud.* **74**, 597–624 (2007).
- Gibrat, R. *Les inégalités économiques* (Librairie du Recueil Sirey, 1931).
- Gabaix, X. Zipf’s law for cities: an explanation. *Q. J. Econ.* **114**, 739–767 (1999).
- Bouchaud, J.-P. Crises and collective socio-economic phenomena: simple models and challenges. *J. Stat. Phys.* **151**, 567–606 (2013).
- Bettencourt, L. & West, G. A unified theory of urban living. *Nature* **467**, 912–913 (2010).
- Soo, K. T. Zipf’s law for cities: a cross-country investigation. *Reg. Sci. Urban Econ.* **35**, 239–263 (2005).
- Singer, H. The “Courbe des populations”. A parallel to Pareto’s law. *Econ. J.* **46**, 254–263 (1936).
- Gabaix, X., Lasry, J. M., Lions, P. L. & Moll, B. The dynamics of inequality. *Econometrica* **84**, 2071–2111 (2016).
- Pumain, D. & Moriconi-Ebrard, F. City size distributions and metropolisation. *GeoJournal* **43**, 307–314 (1997).
- Barthelemy, M. *The Structure and Dynamics of Cities* (Cambridge Univ. Press, 2016).
- Corominas-Murtra, B. & Solé, R. Universality of Zipf’s law. *Phys. Rev. E* **82**, 011102 (2010).
- Pumain, D. Une théorie géographique pour la loi de Zipf. *Reg. Dév.* **36**, 125–150 (2012).
- Marsili, M. & Zhang, Y.-C. Interacting individuals leading to Zipf’s law. *Phys. Rev. Lett.* **80**, 2741–2744 (1998).
- Cottineau, C. MetaZipf. A dynamic meta-analysis of city size distributions. *PLoS One* **12**, e0183919 (2017).
- Benguigui, L. & Blumenfeld-Lieberthal, E. A dynamic model for city size distribution beyond Zipf’s law. *Physica A* **384**, 613–627 (2007).
- Blank, A. & Solomon, S. Power laws in cities population, financial markets and internet sites (scaling in systems with a variable number of components). *Physica A* **287**, 279–288 (2000).
- Krugman, P. Confronting the mystery of urban hierarchy. *J. Jpn. Int. Econ.* **10**, 399–418 (1996).
- Ioannides, Y. & Overman, H. Zipf’s law for cities: an empirical examination. *Reg. Sci. Urban Econ.* **33**, 127–137 (2003).
- Favaro, J. M. & Pumain, D. Gibrat revisited: an urban growth model incorporating spatial interaction and innovation cycles. *Geogr. Anal.* **43**, 261–286 (2011).
- Zanette, D. H. & Manrubia, S. C. Role of intermittency in urban development: a model of large-scale city formation. *Phys. Rev. Lett.* **79**, 523–526 (1997).
- Cottineau, C., Reuillon, R., Chapron, P., Rey-Coyrehourcq, S. & Pumain, D. A modular modelling framework for hypotheses testing in the simulation of urbanisation. *Systems* **3**, 348–377 (2015).
- Bettencourt, L. & Zünd, D. Demography, symmetry and the emergence of universal patterns in urban systems. *Nat. Commun.* **11**, 19 (2020).
- Black, D. & Henderson, V. Urban evolution in the USA. *J. Econ. Geogr.* **3**, 343–372 (2003).
- Eeckhout, J. Gibrat’s law for (all) cities. *Am. Econ. Rev.* **94**, 1429–1451 (2004).
- Benguigui, L. & Blumenfeld-Lieberthal, E. The end of a paradigm: is Zipf’s law universal? *J. Geogr. Syst.* **13**, 87–100 (2011).
- Bouchaud, J. P. & Mézard, M. Wealth condensation in a simple model of economy. *Physica A* **282**, 536–545 (2000).
- Erlander, S. & Stewart, N. F. *The Gravity Model in Transportation Analysis: Theory and Extensions* (VSP, 1990).
- Simini, F., Gonzalez, M. C., Maritan, A. & Barabasi, A.-L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).
- Gnedenko, B. V. & Kolmogorov, A. N. *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, 1954).
- van Kampen, N. G. Itô versus Stratonovich. *J. Stat. Phys.* **24**, 175–187 (1981).
- Srokowski, T. Multiplicative Lévy processes: Itô versus Stratonovich interpretation. *Phys. Rev. E* **80**, 051113 (2009).
- Jespersen, S. M. R. & Fogedby, H. C. Lévy flights in external force fields: Langevin and fractional Fokker–Planck equations and their solutions. *Phys. Rev. E* **59**, 2736–2745 (1999).
- Schertzer, D., Larchevêque, M., Duan, J., Yanovsky, V. & Lovejoy, S. Fractional Fokker–Planck equation for nonlinear stochastic differential equations driven by non-Gaussian Lévy stable noises. *J. Math. Phys.* **42**, 200–212 (2001).
- Fox, C. The G and H functions as symmetrical Fourier kernels. *Trans. Am. Math. Soc.* **98**, 395–429 (1961).
- Clauset, A., Shalizi, C. R. & Newman, M. E. Power-law distributions in empirical data. *SIAM Rev.* **51**, 661–703 (2009).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Article

Methods

For each of the four countries we build a graph of migration flows between metropolitan areas. We have (1) the populations of metropolitan areas and (2) the migration flows between metropolitan areas (described in more detail below).

US migrations

Data of migrations in the USA are taken from the 2013–2017 American Community Survey (ACS)⁴³. Aggregated metro-area-to-metro-area migration flows and counterflows are directly given between 389 metropolitan statistical areas in the USA. More precisely, the ACS asked respondents whether they lived in the same residence one year ago; for people who lived in a different residence, the location of their previous residence was collected.

French interurban migrations

Data of migrations in France are taken from the 2008 INSEE report for residential migrations at the town (commune) level for each individual household⁴⁴. The main residence in 2008 is compared to the main residence in 2003. In order to work at the urban area level, we used the 1999 INSEE list of urban areas and aggregate residential migrations at the metropolitan level, enabling us to analyse migration flows between the 500 largest urban areas in France.

UK interurban migrations

Data of migrations in the UK are taken from 2012–2016 ONS reports on internal migration between English and Welsh local authorities, giving the square matrix of moves each year⁴⁵. In order to work at the urban area level, we used the list of local authorities by OECD functional urban areas and aggregate residential migrations at the metropolitan level, enabling us to analyse migration flows between the 41 largest urban areas in England and Wales.

Canadian interurban migrations

Data of migrations in Canada are taken from 2012–2016 census reports on internal migration between Canadian metropolitan areas⁴⁶. Flows

between these areas are given city-to-city for each year between 2012 and 2016 for the top-160 largest cities in Canada.

Data availability

The datasets used in this study are freely available from public repositories^{43–46}.

43. United States Census Bureau. Metro area-to-metro area migration flows: 2013–2017 American Community Survey; <https://www.census.gov/data/tables/2017/demo/geographic-mobility/metro-to-metro-migration.html> (United States Census Bureau, 2019).
44. INSEE. Migrations résidentielles en 2008: lieu de résidence actuelle – lieu de résidence antérieure; <https://www.insee.fr/fr/statistiques/2022291> (INSEE, 2011).
45. Park, N. Internal migration: matrices of moves by local authority and region (countries of the UK) 2012–2016; <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/migrationwithineuk/datasets/matricesofinternalmigrationmovesbetweenlocalauthoritiesandregionsincludingthecountriesofwalesscotlandandnorthernireland> (Office for National Statistics, 2019).
46. Statistics Canada. Interprovincial and intraprovincial migrants, by census metropolitan area of origin and destination for the period from July 1 to June 30; <https://doi.org/10.25318/1710008701-eng> (Statistics Canada, 2020).

Acknowledgements V.V. is supported by the Ecole nationale des ponts et chaussées and by the Complex Systems Institute of Paris Île-de-France (ISC-PIF). V.V. thanks J. Morán for discussions and comments.

Author contributions V.V. and M.B. designed the study, V.V. acquired the data, and V.V. and M.B. analysed and interpreted the data and wrote the manuscript.

Competing interests The authors declare no competing interests.

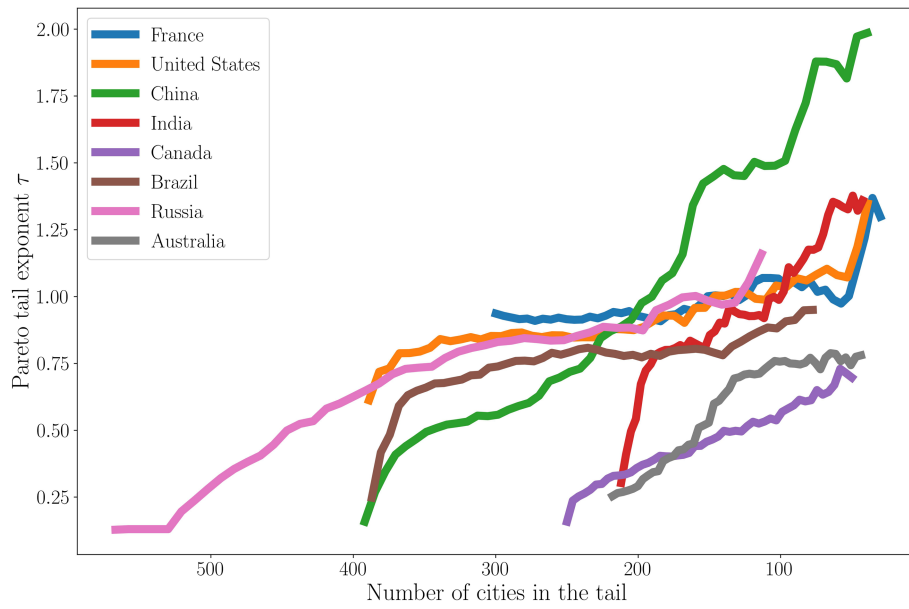
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-2900-x>.

Correspondence and requests for materials should be addressed to M.B.

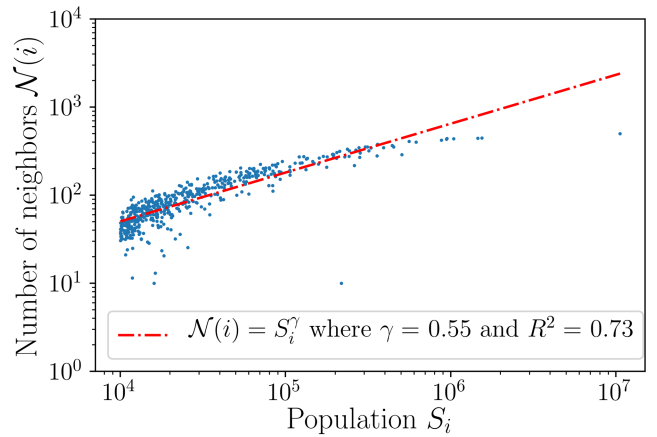
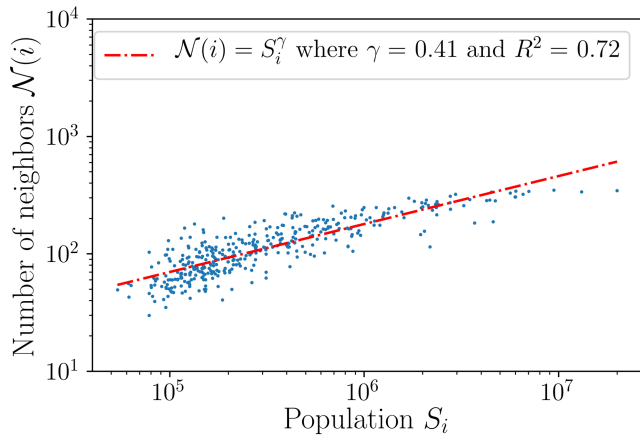
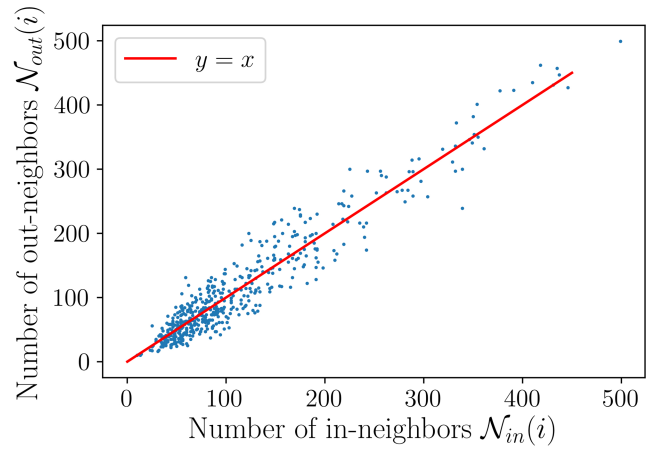
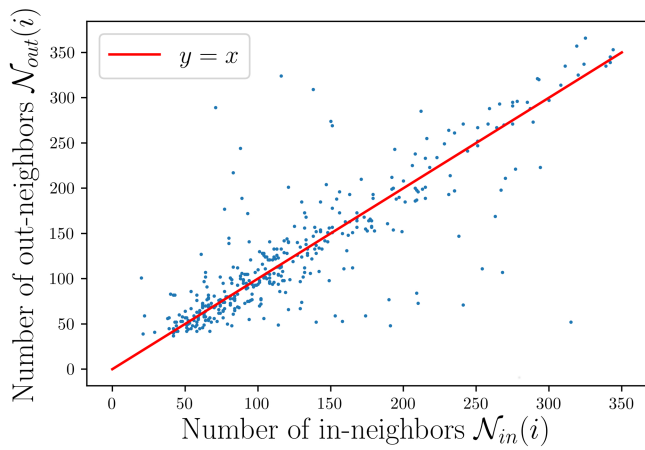
Peer review information *Nature* thanks Amos Maritan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



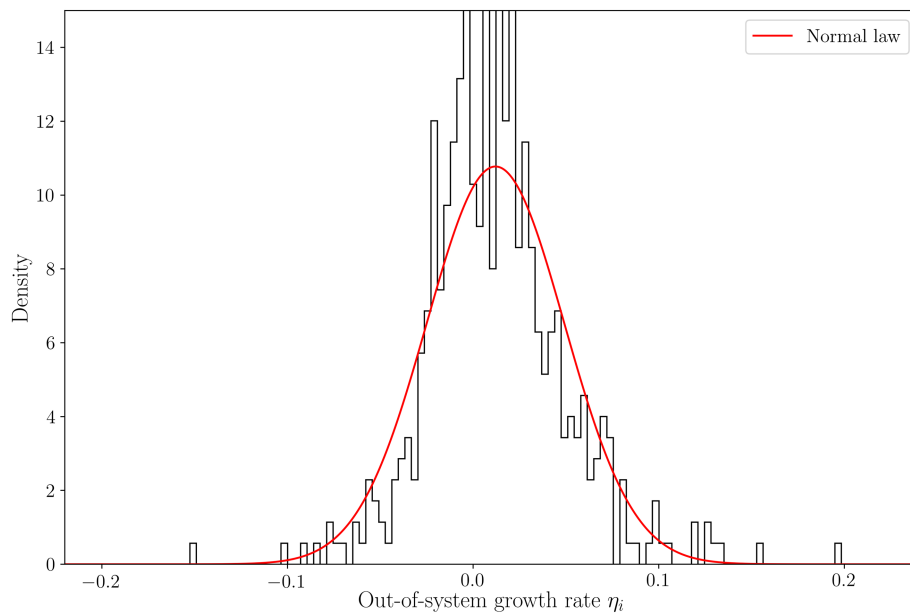
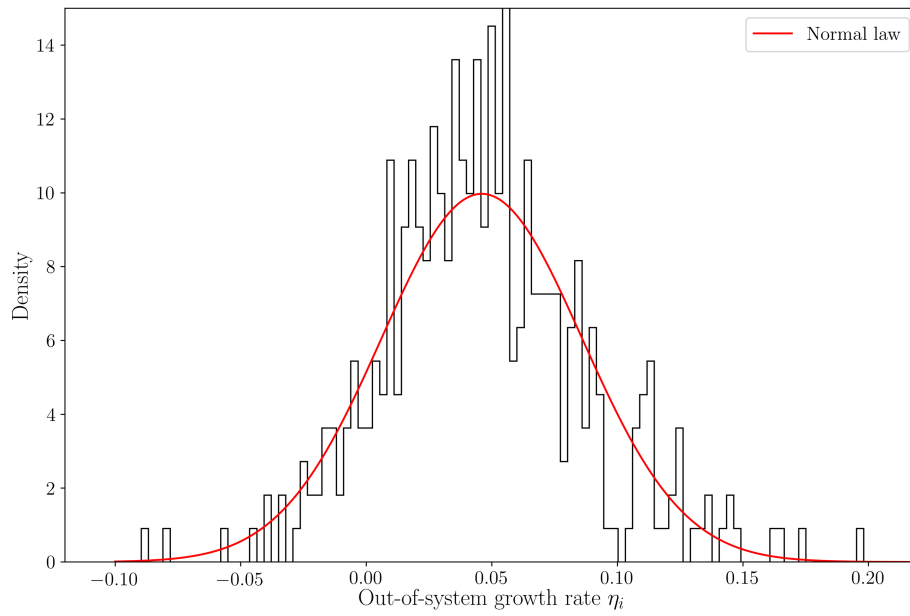
Extended Data Fig. 1 | No universal exponent. We show here the measured Pareto exponent of the upper tail of city-size distributions as a function of the lower threshold defining the tail for the largest cities of eight different

countries. The exponents are obtained with a maximum-likelihood estimate (data from <https://simplemaps.com/data/world-cities>).



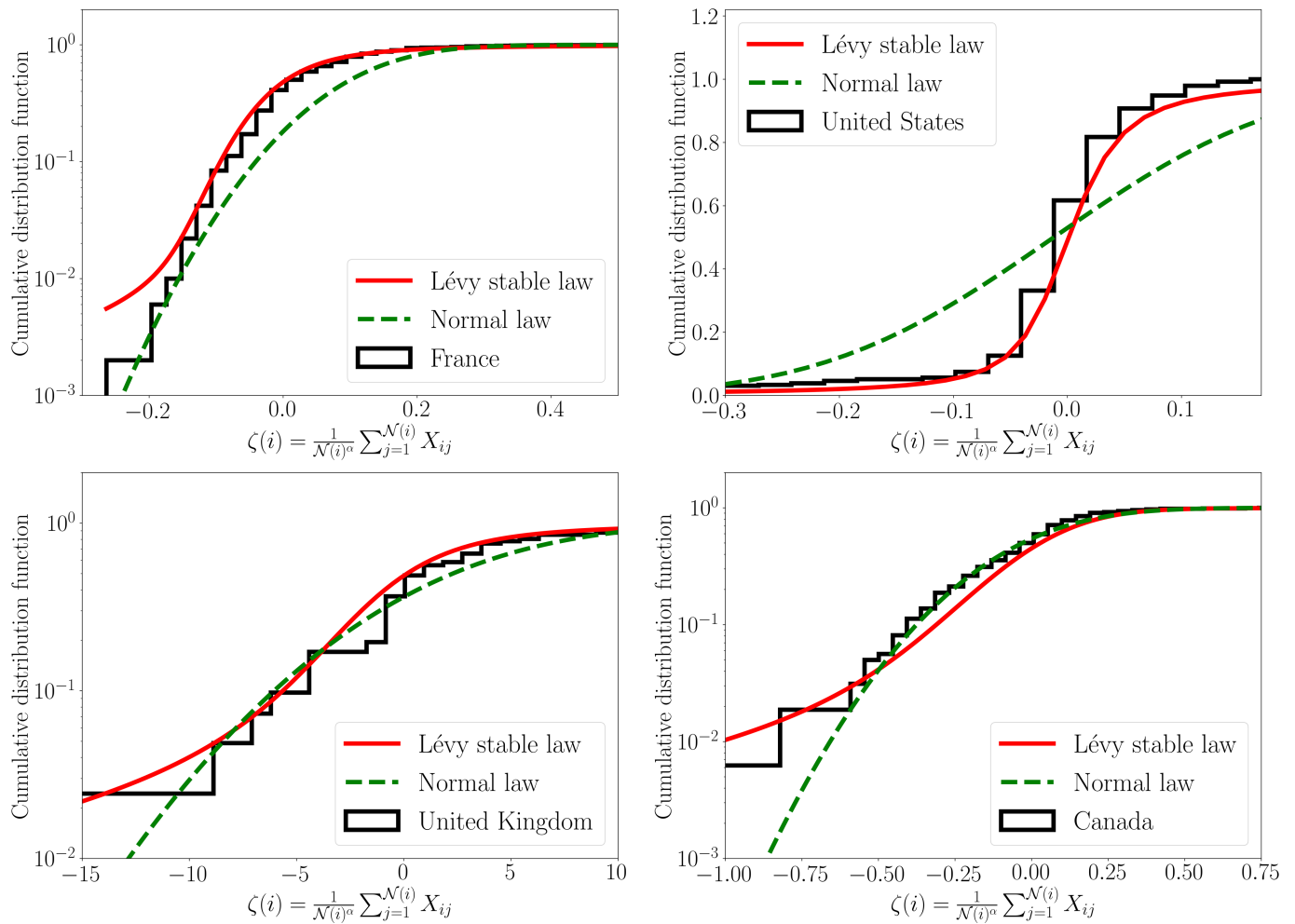
Extended Data Fig. 2 | In- and out-neighbours. a, b, Number of in- and out-neighbours (in the sense of graph theory) for the USA (a) and France (b). The red lines correspond to the equality $N_{in} = N_{out}$. In the UK and Canada, we have a fully connected dataset and $N_{in} = N_{out} = \text{constant}$. **c, d,** Number of

neighbours for each city as a function of population for the USA (c) and France (d). The dotted red lines indicate the power-law fit $|N_i| \propto S_i^\gamma$. In the UK and Canada, we have a variance of the normalized quantity fully connected dataset and $\gamma = 0$.



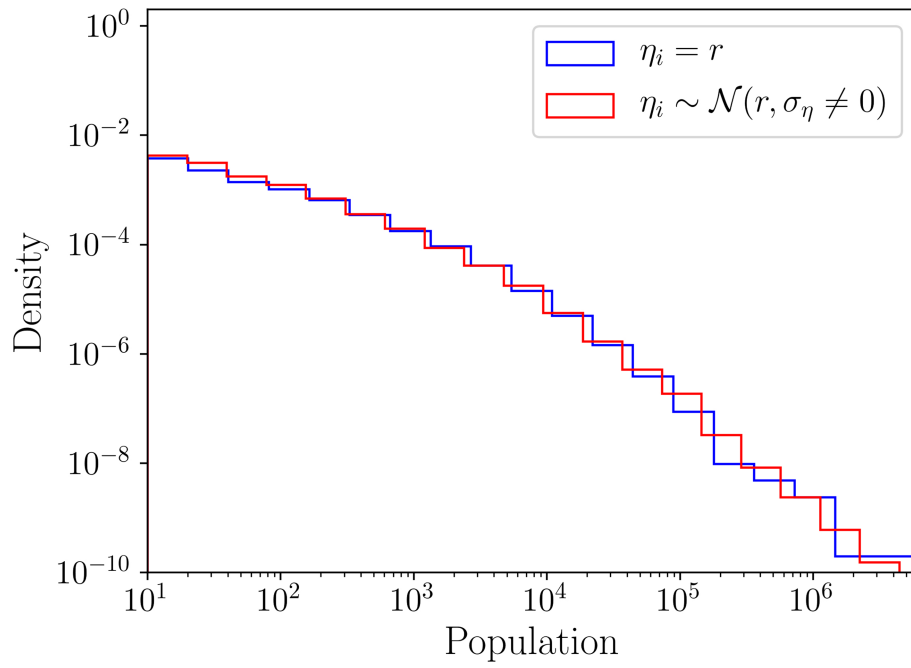
Extended Data Fig. 3 | Density function of the out-of-system growth rate. Natural growth and out-of-system migrations include international migrations and exchanges with small towns. The data shown are for US cities in 2013–2017 (top) and French cities (bottom) in 2003–2008, compared to a normal

distribution. We note that a power-law fit of the right or the left tail would lead to a Pareto exponent of $\beta \gg 1$. For French cities, we extrapolated the 2003 population in each city from the 1999 and 2006 censuses to test our assumption on the period 2003–2008.

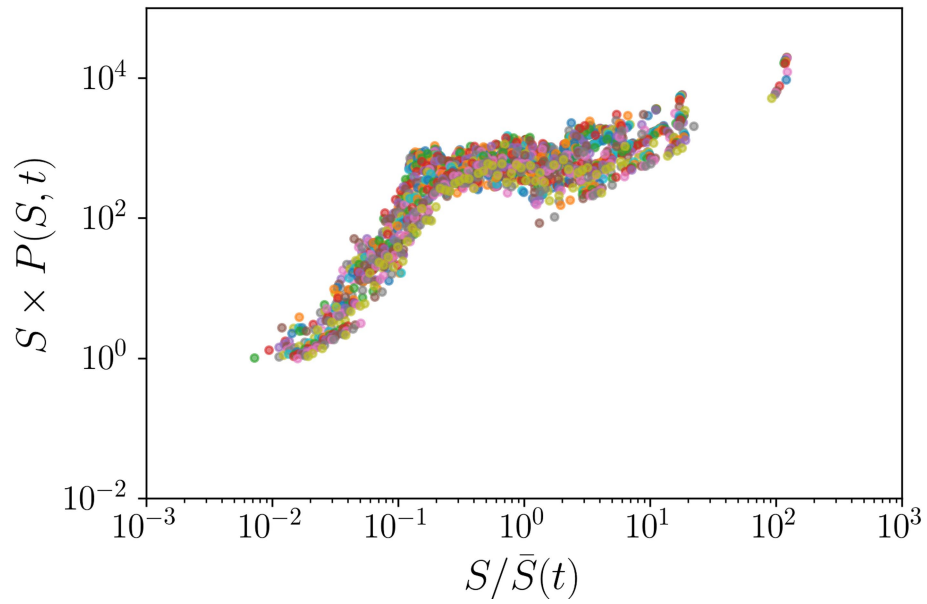


Extended Data Fig. 4 | Migration-flow analysis. Complementary figure to Fig. 2. Empirical left-cumulative distribution function of renormalized migration flows compared to Lévy (continuous red line) and normal

(green dashed lines) distributions. Clockwise from top left, distributions are given for France, the USA, Canada and the UK.

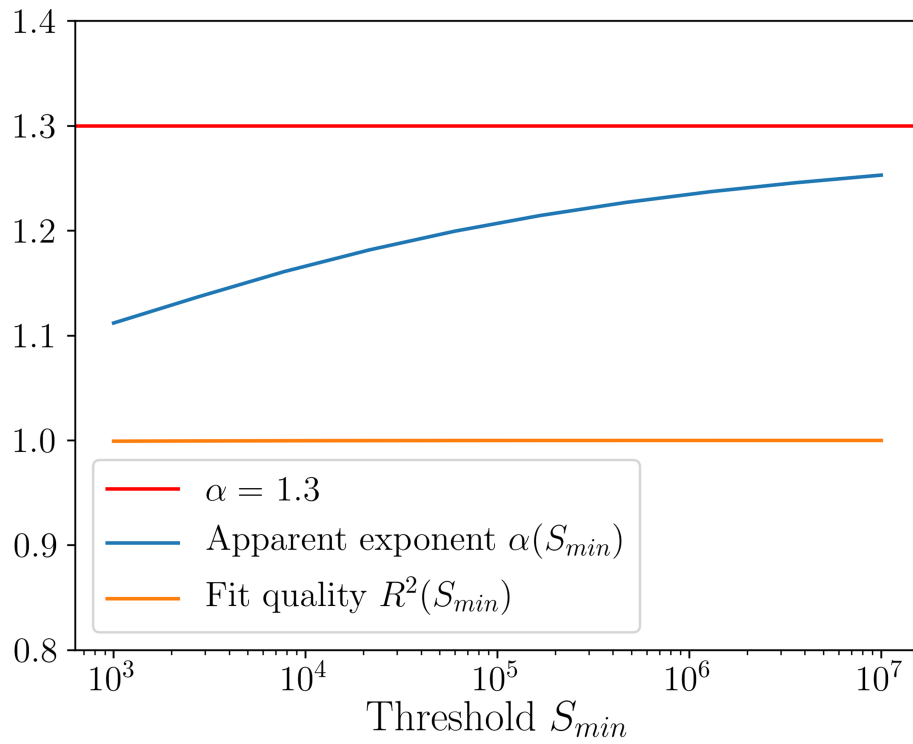


Extended Data Fig. 5 | Average distribution of city sizes. Data obtained by 10 numerical runs of the stochastic differential equation (10) with a Gaussian noise with finite variance σ_η , compared with the numerical solution of equation (10) where $\eta = \langle \eta \rangle = r$. Parameters here are $\alpha = 1.3, \beta = 0.8, r = 0.01, \sigma_\eta = 0.06, D = 0.06$ and $t = 500$.



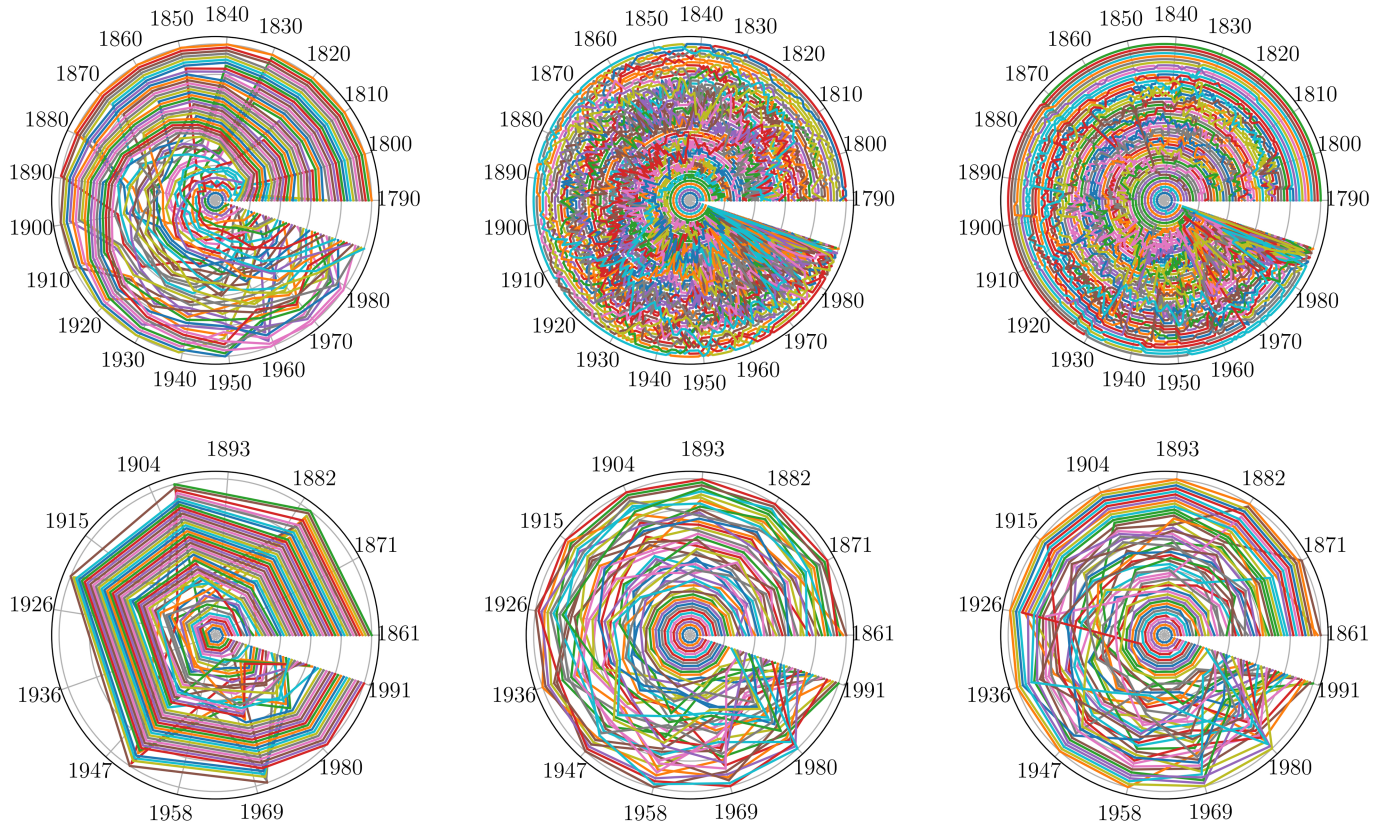
Extended Data Fig. 6 | Scatterplot of the quantity $P(S, t) \times S$ versus the ratio for France's top-500 largest cities between 1875 and 2016. Each colour is a different year. We observe that the plots of all years collapse towards a unique

universal function of the ratio in agreement with the result of equation (33) in Supplementary Information.



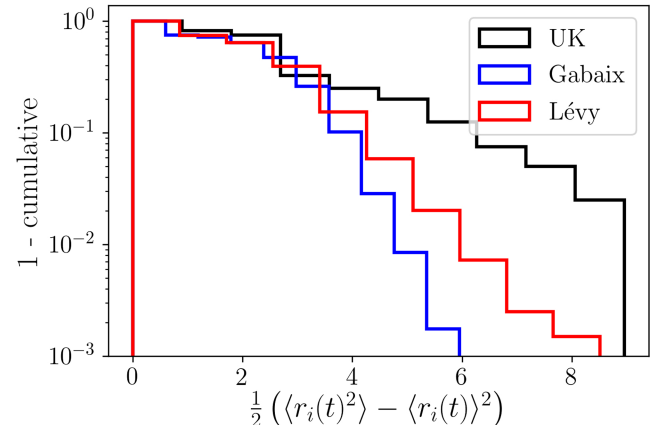
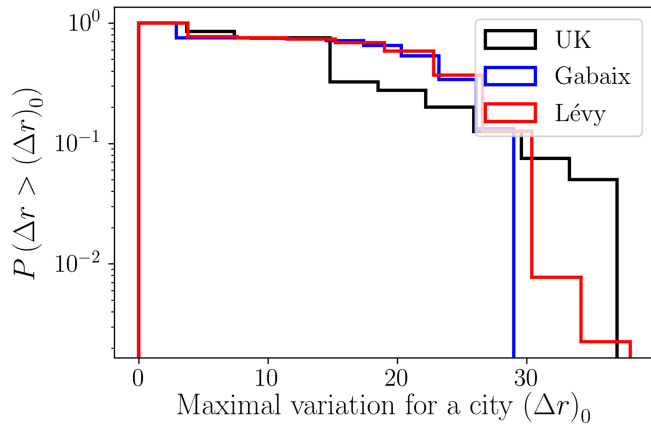
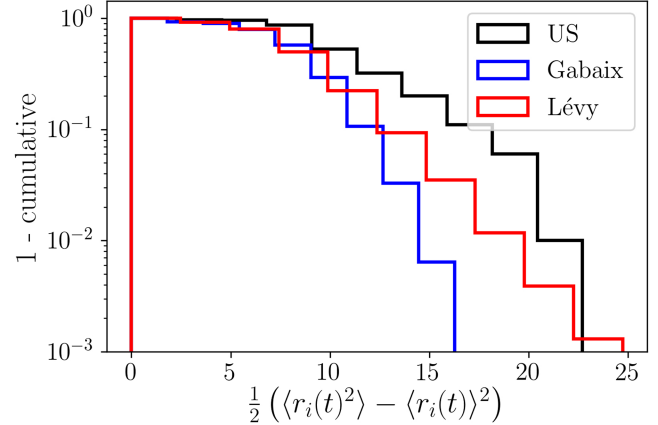
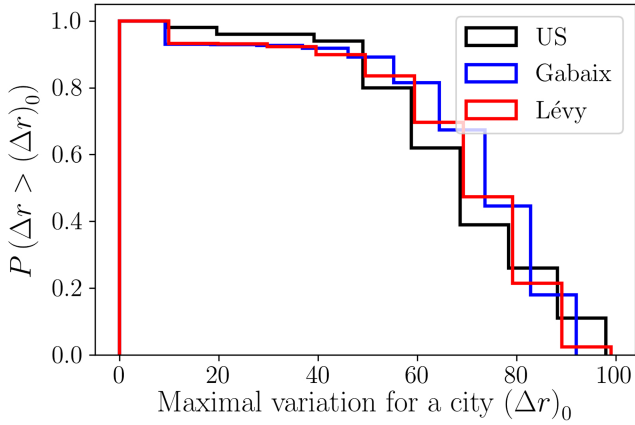
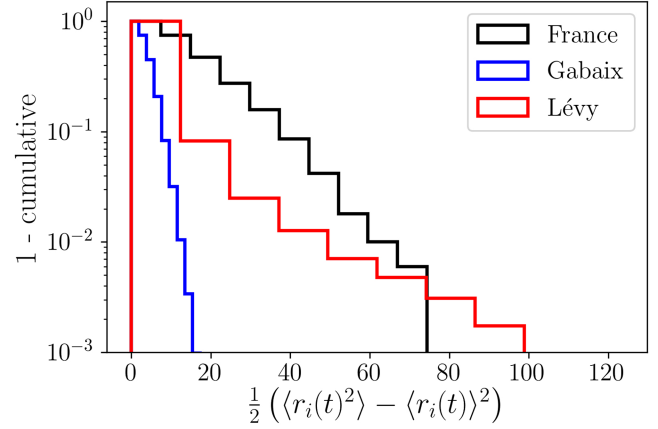
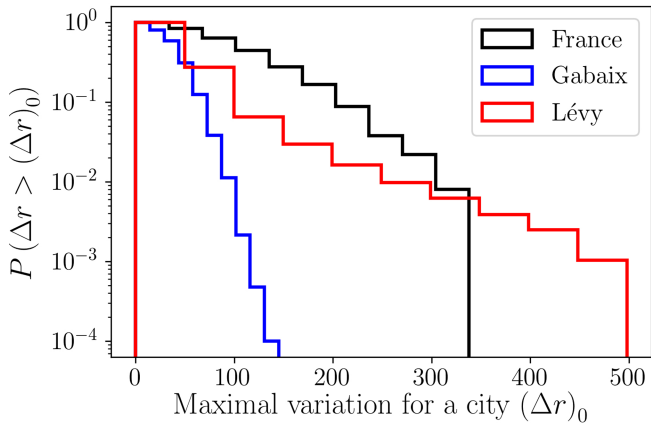
Extended Data Fig. 7 | Power-law fit of the expansion with $\alpha=1.3$, as a function of the lower threshold of city sizes, S_{min} . The expansion is described in equation (25) in Supplementary Information. The fit gives an apparent exponent $\alpha(S_{min})$ with very good quality ($R^2 \approx 1$), although the expansion itself is

not a power law. The apparent exponent is smaller than α , but slowly converges towards $\alpha=1.3$ as the value of the threshold S_{min} increases. Parameters here are $\alpha=1.3, \beta=0.8, r=0.01, D=0.06$ and $t=500$.



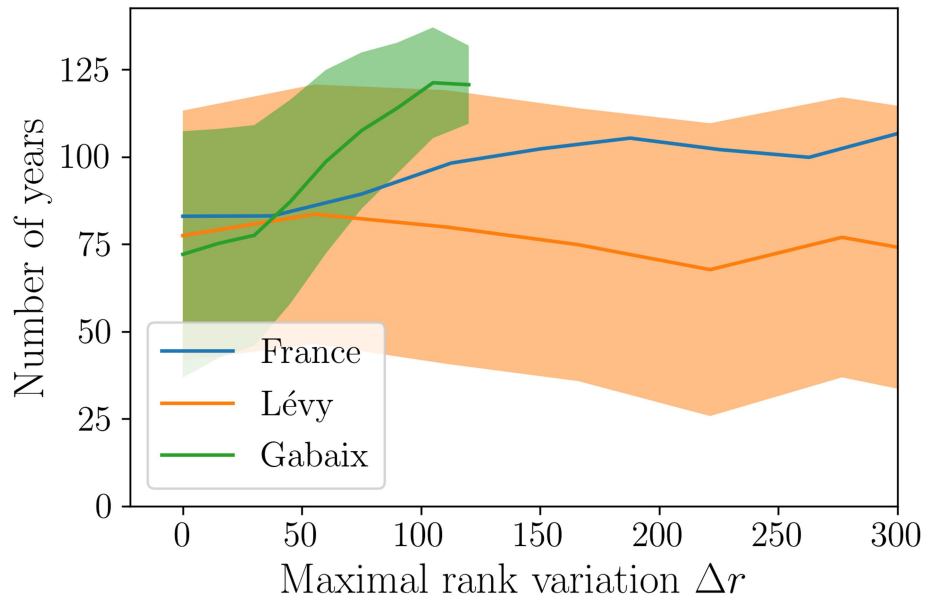
Extended Data Fig. 8 | Rank clocks of the USA and the UK. Top, USA; bottom, UK. The left panels display real data, the middle panels show Gabaix's model of growth and the right panels give our model of growth. Parallel lines for earlier

years are spurious effects resulting from the absence of data for cities out of the top-100 largest in the USA or the top-40 largest in the UK (for these rank clocks, we assigned a random increasing radius to cities without data).



Extended Data Fig. 9 | Microdynamics of city rank through time for the largest cities in France, the USA and the UK. Data is given for the 500 largest cities in France between 1875 and 2016 (top), the 100 largest cities in the United States between 1790 and 1990 (middle) and the 40 largest cities in the UK between 1861 and 1991 (bottom). The left panels display the right-cumulative distribution of the maximal variation of the rank $r_i(t)$ for city, that is, the

difference between the highest and the lowest rank in population for each city. The right panels display the typical fluctuations of the rank $r_i(t)$ through time. In the three cases, the Lévy model is able to predict rare but non-negligible large variations of rank such as the sudden birth or death of city, in contrast to Gabaix's model or Gibrat's model for growth, for which large fluctuations of rank order do not occur.



Extended Data Fig. 10 | Average number of years (and standard dispersion) taken to observe the maximal rank variation Δr as a function of Δr . Although the dispersion is large, Lévy's model is compatible with real data, in contrast to Gabaix's model of growth.