

# Sample and population exponents of generalized Taylor's law

Andrea Giometto<sup>a,b,1</sup>, Marco Formentin<sup>c,d,1</sup>, Andrea Rinaldo<sup>a,e,1</sup>, Joel E. Cohen<sup>f</sup>, and Amos Maritan<sup>c</sup>

<sup>a</sup>Laboratory of Ecohydrology, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; <sup>b</sup>Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, CH-8600 Dübendorf, Switzerland; <sup>c</sup>Dipartimento di Fisica ed Astronomia, Università di Padova, I-35131 Padova, Italy; <sup>d</sup>Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, CZ-18208 Prague, Czech Republic; <sup>e</sup>Dipartimento di Ingegneria Civile, Edile ed Ambientale, Università di Padova, I-35131 Padova, Italy; and <sup>f</sup>Laboratory of Populations, The Rockefeller University and Columbia University, New York, NY 10065-6399

Contributed by Andrea Rinaldo, March 27, 2015 (sent for review January 14, 2015; reviewed by Pablo A. Marquet)

**Taylor's law (TL) states that the variance  $V$  of a nonnegative random variable is a power function of its mean  $M$ ; i.e.,  $V = aM^b$ . TL has been verified extensively in ecology, where it applies to population abundance, physics, and other natural sciences. Its ubiquitous empirical verification suggests a context-independent mechanism. Sample exponents  $b$  measured empirically via the scaling of sample mean and variance typically cluster around the value  $b = 2$ . Some theoretical models of population growth, however, predict a broad range of values for the population exponent  $b$  pertaining to the mean and variance of population density, depending on details of the growth process. Is the widely reported sample exponent  $b \simeq 2$  the result of ecological processes or could it be a statistical artifact? Here, we apply large deviations theory and finite-sample arguments to show exactly that in a broad class of growth models the sample exponent is  $b \simeq 2$  regardless of the underlying population exponent. We derive a generalized TL in terms of sample and population exponents  $b_{jk}$  for the scaling of the  $k$ th vs. the  $j$ th cumulants. The sample exponent  $b_{jk}$  depends predictably on the number of samples and for finite samples we obtain  $b_{jk} \simeq k/j$  asymptotically in time, a prediction that we verify in two empirical examples. Thus, the sample exponent  $b \simeq 2$  may indeed be a statistical artifact and not dependent on population dynamics under conditions that we specify exactly. Given the broad class of models investigated, our results apply to many fields where TL is used although inadequately understood.**

fluctuation scaling | multiplicative growth | power law | environmental stochasticity | Markovian environment

**T**aylor's law (TL) (1), also known as fluctuation scaling in physics, is one of the most verified patterns in both the biological (2–6) and physical (7–12) sciences. TL states that the variance of a nonnegative random variable  $V = \text{Var}[X]$  is approximately related to its mean  $M = \mathbb{E}[X]$  by a power law; that is,  $\text{Var}[X] = a\mathbb{E}[X]^b$ , with  $a > 0$  and  $b \in \mathbb{R}$ . In ecology, the random variable of interest is generally the size or density  $N$  of a censused population and TL can arise in time (i.e., the statistics of  $N$  are computed over time) or in space (i.e., the statistics are computed over space). The widespread verification of TL has led many authors to suggest the existence of a universal mechanism for its emergence, although there is currently no consensus on what such a mechanism would be. Various approaches have been used in the attempt, ranging from the study of probability distributions compatible with the law (13–15) to phenomenological and mechanistic models (16–20). Although most empirical studies on spatial TL report an observed sample exponent  $b$  in the range 1–2 (1, 21), mostly around  $b \simeq 2$  (21) [figure 10(g) in ref. 22], population growth models (5, 23–26) can generate TL with any real value of the exponent. Moreover, theoretical investigations of multiplicative growth models in correlated Markovian environments (24, 25) have shown that the exponent  $b$  can undergo abrupt transitions following smooth changes in the environmental autocorrelation.

Here, we distinguish between values of  $b$  derived from empirical fitting (sample exponents) and values obtained via theoretical models that pertain to the probability distribution of the random variable  $N$  (population exponents). We show that in a broad class of multiplicative growth models, the sample and population exponents coincide only if the number of observed samples or replicates is greater than an exponential function of the duration of observation. Among the relevant consequences, we demonstrate that the sample TL exponent robustly settles on  $b \simeq 2$  for any Markovian environment observed for a duration that is larger than a logarithmic function of the number of replicates. Accordingly, when the number of observations is limited, abrupt transitions in the sample TL exponent can be observed only within relatively short time windows.

## Results

Let us consider multiplicative growth models in Markovian environments (24, 25). Let  $N(t)$  be the density of a population at time  $t$  and assume that the initial density is  $N_0 > 0$ .  $N(t)$  is assumed to undergo a multiplicative growth process such that

$$N(t) = N_0 \prod_{n=1}^t A_n. \quad [1]$$

The values of the multiplicative growth factors  $A_i$  are determined via a two-state homogeneous Markov chain with state space

## Significance

**Taylor's law (TL) has been verified very widely in the natural sciences, information technology, and finance. The widespread observation of TL suggests that a context-independent mechanism may be at work and stimulated the search for processes affecting the scaling of population fluctuations with population abundance. We show that limited sampling may explain why TL is often observed to have exponent  $b = 2$ . Abrupt transitions in the TL exponent associated with smooth changes in the environment were recently discovered theoretically and comparable real-world transitions could harm fish populations, forests, and public health. Our study shows that limited sampling hinders the anticipation of such transitions and provides estimates for the number of samples required to reveal early warning signals of abrupt biotic change.**

Author contributions: A.G., M.F., A.R., J.E.C., and A.M. designed research, performed research, analyzed data, and wrote the paper.

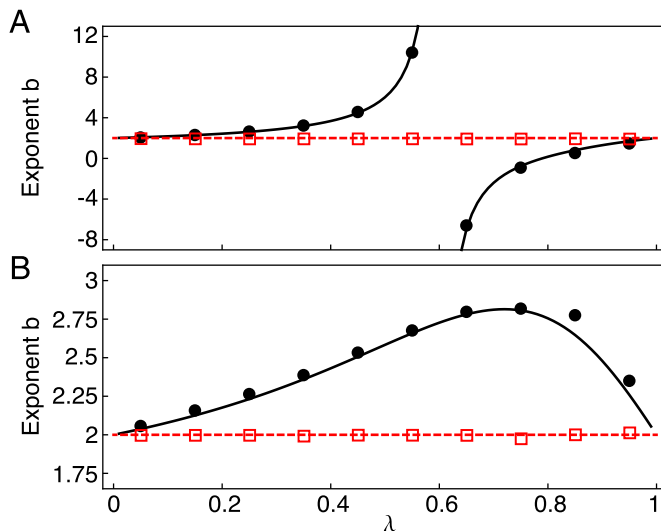
Reviewers included: P.A.M., Catholic University of Chile.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence may be addressed. Email: andrea.rinaldo@epfl.ch, andrea.giometto@epfl.ch, or marco.formentin@ruhr-uni-bochum.de.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1505882112/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1505882112/-DCSupplemental).



**Fig. 1.** TL exponent  $b$  for different values of the transition probability  $\lambda$ . The sample exponents computed in  $R$  simulations of a two-state multiplicative process with symmetric transition matrix in the two regimes  $1 \ll t \ll \log R$  (black solid circles,  $R = 10^5$  up to time  $t = 10$ ) and  $t \gg \log R$  (red open squares,  $R = 10^4$  up to time  $t = 400$ ) are in good agreement with predictions for the asymptotic population (black solid line, Eq. 9) and sample (red dashed line,  $b = 2$ ) exponents. In the simulations, the sample exponent  $b$  was computed by least-squares fitting of  $\log \text{Var}[N(t)]$  as a function of  $\log \mathbb{E}[N(t)]$  for the last 6 (black circles) and 200 (red squares) time steps. In *A*, which has the plotted theoretical result from ref. 24,  $\chi = \{r, s\} = \{2, 1/4\}$  ( $b$  displays a discontinuity); in *B*,  $\chi = \{r, s\} = \{4, 1/2\}$  (in such a case,  $b$  displays no discontinuity). Fig. S1 shows the generalized TL exponent  $b_{23}$  in the same simulations.

$\chi = \{r, s\}$  (we assume, without loss of generality,  $r > s$  and  $N_0 = 1$ ) and transition matrix  $\Pi$  with  $\Pi(i, j) > 0$  for all  $i, j \in \chi$  (Methods). In our notation,  $\Pi(i, j)$  is the one-step probability to go from state  $i$  to state  $j$ ; i.e.,  $\Pi(i, j) = \text{Prob}(A_{n+1} = j | A_n = i)$ . For the sake of clarity, we restrict our discussion to symmetric transition matrices, with  $\Pi(i, j) = \lambda$  for  $i \neq j$ . We derive (Methods) exact results on both sample and population TL exponents for a broad class of multiplicative processes, including state spaces with size higher than 2 and nonsymmetric transition matrices (SI Methods).

By adopting large deviation theory techniques (27, 28) and finite sample size arguments (29), we show (Methods) that for any choice of  $\Pi$  and  $\chi$ , the sample mean and variance in a finite set of  $R$  independent realizations of the process obey TL asymptotically as  $t \rightarrow \infty$  with an exponent that may differ from the corresponding population exponent. More precisely, our analysis reveals two regimes ( $t \gg \log R$  and  $1 \ll t \ll \log R$ , respectively—all logarithms here are to the base  $e$ ) where the sample TL holds with different exponents. In the former regime, sample exponents inevitably tend to  $b \simeq 2$  independently of model specifications. In the latter, sample exponents accurately approximate population ones, which can be computed analytically and may differ from  $b = 2$ . Fig. 1 shows that simulation results and theoretical predictions in the two regimes are in excellent agreement. Fig. 2 shows the temporal evolution of the sample TL exponent, which crosses over from the approximate value of the population exponent (Eq. 9) at small times to the asymptotic prediction  $b \simeq 2$  at larger times (Eq. 13).

We derive a generalized TL that involves the scaling of the  $k$ th moment vs. the  $j$ th moment of the distribution of  $N(t)$ . Exact results (Methods) show that the generalized TL,

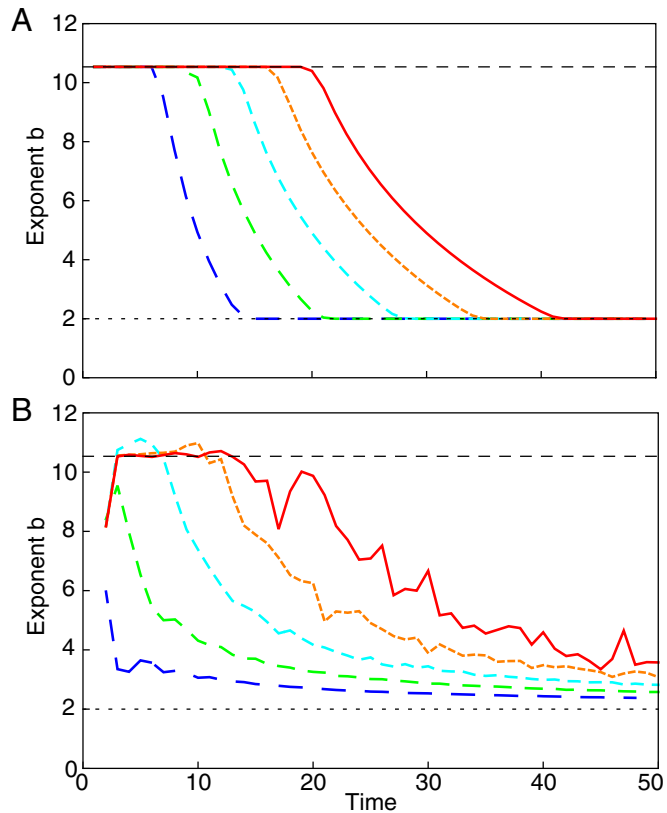
$$\mathbb{E}[N^k(t)] = a_{jk} \mathbb{E}[N^j(t)]^{b_{jk}}, \quad [2]$$

holds asymptotically in  $t$  for any choice of  $j$  and  $k$  (including noninteger values), both for population and for sample moments

(the positivity of  $\Pi$  ensures that the same scaling relationship holds between the  $k$ th and  $j$ th cumulants) (SI Methods). In accordance with the above results on the conventional TL (recovered in this framework with the choice  $j = 1, k = 2$ ), two regimes exist: If  $1 \ll t \ll \log R$ , sample moments and cumulants accurately approximate population ones (and the value of  $b_{jk}$  can be computed analytically); if  $t \gg \log R$ , the generalized TL exponent approximates  $b_{jk} \simeq k/j$  (Fig. S1 C and D).

In ecological contexts, the number of realizations  $R$  that determine the possible convergence of sample and population TL exponents could refer, for instance, to independent patches experiencing different realizations of the same climate (24). In an established ecosystem, species have been present for several generations, and one might assume that the system is in the asymptotic regime  $t \gg \log R$ . Within this perspective, we tested the prediction that for large  $t$  sample exponents satisfy the relation  $b_{jk} = k/j$  (including the conventional TL) on two datasets.

A first example is drawn from a long-term census of plots within the Black Rock Forest (BRF) (5). It was shown that the Lewontin-Cohen model (a particular case of the multiplicative model studied here) describes the population dynamics of trees in BRF and provides an interpretation of the TL exponent (5). The interpretation of the six plots as distinct and independent



**Fig. 2.** Time evolution of the sample TL exponent. The sample exponent {computed as the slope of the curve  $\log \mathbb{E}[N(t)^2]$  vs.  $\log \mathbb{E}[N(t)]$ } crosses over from the approximate population exponent (Eq. 9, dashed upper horizontal line) at small times to  $b \simeq 2$  (dotted lower horizontal line) at larger times. The number of simulations  $R = 10^n$  increases exponentially from  $10^2$  (blue dashed lines) to  $10^6$  (red solid lines), whereas the crossover time increases approximately linearly. Here,  $\chi = \{r, s\} = \{2, 1/4\}$  and the transition probability (with  $\Pi$  symmetric) is  $\lambda = 0.55$ . (A) Theoretical prediction computed via Eq. 15. (B) Simulation results. Curves are averaged over  $10^8/R$  simulations (except for the blue curve, averaged over  $10^5$  simulations). Mismatches between *A* and *B* are due to the necessity to have  $t$  and  $R$  not too large to keep simulations feasible, whereas Eqs. 8 and 13 hold true asymptotically in  $t$ .

replicates of the Lewontin–Cohen model is supported by statistical analysis (5) and allowed relating the model predictions to the spatial TL. Here, we computed, for each year  $t$ , the spatial sample moments  $\langle N^k \rangle(t)$  of tree abundance across plots and we found that the least-squares slopes  $b_{jk}$  of  $\log\langle N^k(t) \rangle$  vs.  $\log\langle N^j(t) \rangle$  (Table 1) are compatible with the asymptotic model prediction  $b_{jk} = k/j$  (Methods).

A second example uses the data collected by P. den Boer (30), who measured abundances of carabid beetles in various sites across The Netherlands within a 200-km<sup>2</sup> area for 8 consecutive years. The dataset was shown to support the conventional spatial TL (16). We computed the sample moments of carabid beetles abundance,  $\langle N^k(t) \rangle$ , across similar sites (either woodland or heath), for each species separately and year  $t$ . In the intraspecific analysis (Fig. 3), linear regressions of  $\log\langle N^k(t) \rangle$  vs.  $\log\langle N^j(t) \rangle$  for  $t = 1, \dots, Y$  ( $Y$  is the total number of years) gave the estimate of the sample exponent  $b_{jk}$  for each species separately (different species are identified by different colors in Fig. 3). Frequency histograms of empirical exponents  $b_{jk}$  are shown in Fig. S2 (also the box-whisker plots in Fig. 3, *Insets*); for every integer choice of  $j$  and  $k$  (here, up to  $k = 4$ ), the histogram is centered in  $k/j$ , as the asymptotic model predicted. A one-sample  $t$  test does not reject the null hypothesis that the sample mean of the values of  $b_{jk}$  does not differ significantly from the theoretically predicted mean  $k/j$  (Fig. S2). In the interspecific analysis (Fig. 4), we calculated the least-squares slope  $b_{jk}$  (for  $j = 1$ ) of  $\log\langle N^k \rangle$  vs.  $\log\langle N \rangle$  across all species at a given year and site type (Tables S1 and S2). Each data point in Fig. 4 refers to a single species. The empirical exponents  $b_{jk}$  for all years are compatible with the asymptotic model prediction  $b_{jk} = k/j$ , as are the mean (across years and site type) exponents  $b_{jk}$  (Table 2).

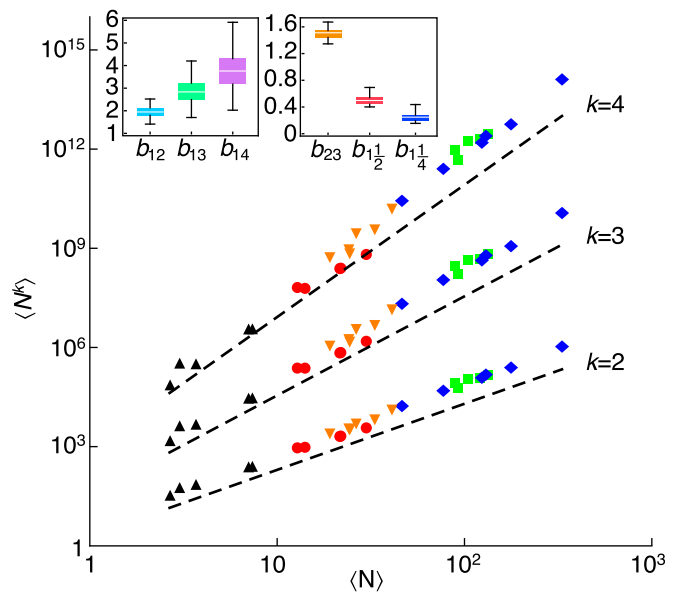
The empirical confirmation and the finding that other demographic models predict the generalized TL with  $b_{jk} = k/j$  (SI Text) indicate that these predictions are probably insensitive to the details of the dynamics, just as the original TL is quite robust (3, 15, 31).

## Discussion

Understanding to what extent widely reported macroecological patterns are the result of statistical instead of ecological processes is one of the main challenges in ecology (32). Here, we have uncovered a general mechanism that yields TL with the widely observed sample exponent  $b \simeq 2$ , which may be attributable to the finite size of both ecosystems and sampling efforts. For a broad range of parameters within the class of multiplicative models, and other demographic processes, the generalized TL describes the scaling of moments and cumulants with the sample exponent  $b_{jk}$  asymptotically equal to  $k/j$ . Our theoretical predictions are supported by two empirical examples and invite further testing, also outside the field of ecology. When the number of samples is limited, TL may not reflect (or depend on) the underlying population dynamics and the empirically measured sample exponent may be a statistical artifact that is

**Table 1. Sample exponents for the generalized TL in the Black Rock Forest dataset, data from ref. 5**

$j, k$	$k/j$	$b_{jk} \pm \text{SE}$	$R^2$
1, 2	2	$2.14 \pm 0.12$	0.991
1, 3	3	$3.33 \pm 0.32$	0.973
1, 4	4	$4.54 \pm 0.58$	0.954
2, 4	2	$2.15 \pm 0.16$	0.984
2, 3	1.5	$1.57 \pm 0.07$	0.995
3, 4	1.333	$1.37 \pm 0.04$	0.997
1, 1/2	0.5	$0.48 \pm 0.02$	0.997
1, 1/4	0.25	$0.23 \pm 0.01$	0.993
1, 2/3	0.667	$0.65 \pm 0.01$	0.999



**Fig. 3. Generalized TL for intraspecific patterns of carabid beetles abundance.** Shown is a double logarithmic plot of  $\langle N^k \rangle$  vs.  $\langle N \rangle$  for different species (identified by different colors and symbols), for consecutive years (each symbol refers to a single year  $t$ ). For visual clarity, only five species are shown. Dashed black lines of slopes  $b_{jk} = k$  (asymptotic model prediction) are shown. Vertical offsets are introduced to aid comparison of slopes. (*Insets*) Box and whisker plots for the empirical distribution of intraspecific generalized TL exponents  $b_{jk}$ , showing the median (white horizontal line) and the 25% and 75% quantiles.

not representative of the population distribution of abundances. Our investigation provides a tool to discern whether the observed patterns of population abundance depend on the underlying population dynamics.

Limited sampling efforts might hinder the observation of abrupt transitions in population exponents that were recently discovered for theoretical multiplicative growth processes. Because fluctuations in population abundances strongly affect ecological dynamics, in particular extinction risk, comparable real-world transitions may harm fish populations, forests, and public health. Our calculation of the minimum number of samples required to observe such transitions may help to identify early signals of abrupt biotic change following smooth changes in the environment.

## Methods

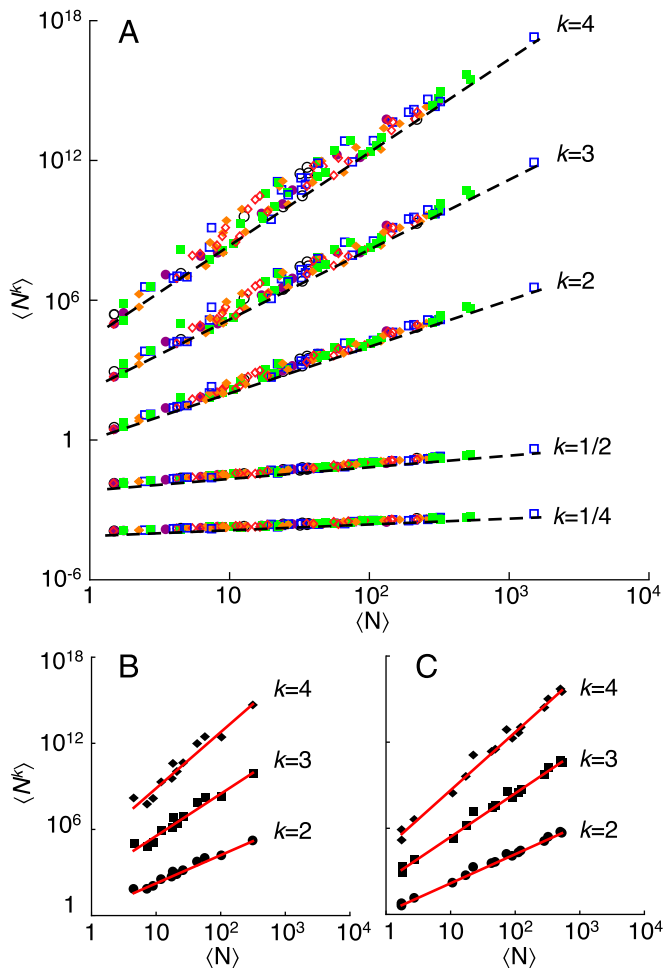
**Theoretical Analysis.** Let  $\Pi$  be a  $2 \times 2$  symmetric matrix. The stationary distribution  $\pi$  of the chain is unique and in the symmetric case satisfies  $\pi(i) = 1/2$ ,  $i \in \mathcal{X}$ , for all  $\lambda \in (0, 1)$ . We assume that the chain starts at equilibrium. We introduce the empirical mean  $L_t(z) : \mathcal{X} \rightarrow [0, 1]$ , defined as

$$L_t(z) = \frac{1}{t} \sum_{n=1}^t \delta_{A_n, z}, \quad [3]$$

where  $\delta$  is Kronecker's delta. The random measure  $L_t(r)$  gives the fraction of times that  $r$  appears in a realization of the Markov chain up to time  $t$ .  $L_t$  satisfies a large deviation principle (LDP) (27) with rate function

$$I_\Pi(x) = \sup_{u>0} \left[ x \log \left( \frac{u_1}{(\Pi u)_1} \right) + (1-x) \log \left( \frac{u_2}{(\Pi u)_2} \right) \right], \quad [4]$$

where  $x$  ( $x \in [0, 1]$ ) is the proportion of  $r$  in a realization of the Markov chain up to time  $t$  (correspondingly, the proportion of  $s$  is  $1-x$ ) and  $u$  is a strictly positive vector in  $\mathbb{R}^2$  (i.e.,  $u_1, u_2 > 0$ ). Stating that  $L_t$  satisfies a LDP means that  $\lim_{t \rightarrow \infty} (1/t) \log \mathbb{P}(L_t(r) \in [x, x+dx]) = -I_\Pi(x)$ . The rate function  $I_\Pi(x)$  is convex ( $d^2 I_\Pi / dx^2 > 0$ ), attains its minimum at  $x_{\min} = 1/2$  with  $I_\Pi(x_{\min}) = 0$ , and is symmetric around  $x_{\min}$  (lemma IV.10 of ref. 27, theorems 3.1.2 and 3.1.6 of



**Fig. 4.** Generalized TL for interspecific patterns of abundance of carabid beetles. (A) Double logarithmic plot of  $\langle N^k \rangle$  vs.  $\langle N \rangle$  for all species, years, and site types. Each data point refers to a single species in 1 y and site type. The color and symbol code identifies data relative to the same year: 1961 (black open circles), 1962 (purple solid circles), 1963 (blue open squares), 1964 (green solid squares), 1965 (orange solid diamonds), and 1966 (red open diamonds). Dashed black lines of slope  $b_{1k} = k$  (asymptotic model prediction) are plotted next to the corresponding data series. Vertical offsets are introduced to aid comparison of slopes. (B and C) Examples of interspecific moments scaling (each data point refers to a single species) for a single year and site type (B, woodland 1964; C, heath 1964) used for the statistical analysis (SI Methods, Table 2, Tables S1 and S2, and Fig. S5). The red lines are the least-squares regressions of  $\log\langle N^k \rangle$  vs.  $\log\langle N \rangle$  across species.

ref. 33, and section 4.3 of ref. 28). The subscript  $\Pi$  is used to indicate that the rate function depends on the transition matrix. Additionally, Eq. 4 depends on  $u_1$  and  $u_2$  only through  $u \equiv u_2/u_1$ ; thus, by standard one-variable calculus, a long but explicit form of  $I_{\Pi}(x)$  can be computed,

$$I_{\Pi}(x) = (x-1) \log \left\{ 1 - \lambda \left[ \frac{2(\lambda-1)x}{S_{\lambda}(x) - 2\lambda x} + 1 \right] \right\} - x \log \left[ 1 - \frac{\lambda(S_{\lambda}(x) - 2x)}{2(\lambda-1)x} \right], \quad [5]$$

where

$$S_{\lambda}(x) = \lambda + \sqrt{\lambda^2 + 8\lambda(x-1)x - 4(x-1)x}. \quad [6]$$

The rate function does not depend on the values of the multiplicative factors  $r$  and  $s$ . As in ref. 25, we consider the ratio between  $t^{-1} \log \text{Var}[N(t)]$  and  $t^{-1} \log \mathbb{E}[N(t)]$ , but here we exploit the LDP, adopting Varadhan's lemma (theorem III.13 of ref. 27), to perform such computation. First, because  $\Pi$  is positive and  $r \neq s$ , it holds true that

$$\lim_{t \rightarrow \infty} t^{-1} \log \text{Var}[N(t)] = \lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^2]. \quad [7]$$

See the appendix in ref. 25 for a proof. Then, for the population moments of the population density  $N(t)$ , applying Varadhan's lemma, we have

$$\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^k] = \sup_{x \in [0,1]} [kG(x) - I_{\Pi}(x)], \quad [8]$$

where  $G(x) = x \log r + (1-x) \log s$ . The population TL exponent  $b$  (which depends on  $\lambda$ ) can thus be computed as

$$b(\lambda) = \frac{\sup_{x \in [0,1]} [2G(x) - I_{\Pi}(x)]}{\sup_{x \in [0,1]} [G(x) - I_{\Pi}(x)]}. \quad [9]$$

For certain values of  $r$  and  $s$ ,  $b(\lambda)$  can show a discontinuity at a critical value of the transition probability  $\lambda$  (black line in Fig. 1A). The existence of such discontinuity was discovered and discussed in ref. 24. An analysis of the critical transition probability is available in SI Methods (Figs. S3 and S4). A generalized TL can be derived by adapting Eq. 8 to compute the scaling exponent for any pair of population moments as

$$b_{jk}(\lambda) = \frac{\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^k]}{\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^j]} = \frac{\sup_{x \in [0,1]} [kG(x) - I_{\Pi}(x)]}{\sup_{x \in [0,1]} [jG(x) - I_{\Pi}(x)]}. \quad [10]$$

Discontinuities can also arise for these population exponents (SI Methods).

Eqs. 9 and 10 hold true when one considers an infinite number of realizations of the multiplicative process, which ensures visiting the whole region  $x \in [0,1]$ . We now estimate the sample exponent  $b$  that is based on the sample mean and variance calculated over a finite set of  $R$  realizations of the multiplicative process. We present here a heuristic derivation of the sample exponent. A more rigorous calculation is given in SI Methods. We define  $x_+$  as the value in  $[0,1]$  such that the probability of a larger frequency  $x$  of  $r$  in  $R$  runs of the Markov chain up to time  $t$  is  $1/R$ :

$$\mathbb{P}(L_t(r) \in (x_+, 1]) = \frac{1}{R}. \quad [11]$$

With this definition,  $x_+$  can be interpreted (29) as the typical maximum frequency of  $r$  in  $R$  realizations of the chain. Analogously, we define  $x_-$  as the value such that smaller values of the frequency of  $r$  are observed with probability  $1/R$ , namely  $\mathbb{P}(L_t(r) \in [0, x_-]) = 1/R$ . For large  $t$ , one can adapt Varadhan's lemma (or Laplace's method of integration) to obtain, as a function of  $t$ , the approximate number of replicas  $R$  needed to explore rare events [i.e., to compute  $\mathbb{P}(L_t(r) \in (x_+, 1]) = R^{-1}$ ]. Approximately

$$R \approx \exp[tI_{\Pi}(x_{\pm})]. \quad [12]$$

Inversion of this formula (by taking the logarithm on both sides and expanding  $I_{\Pi}$  in Taylor series around  $x = x_{\min}$ ) gives  $x_{\pm} \approx 1/2 \pm \sqrt{((1-\lambda)/2\lambda)((\log R)/t)}$ . Consequently, the sample TL exponent in an ensemble of  $R$  realizations of the process can be approximated as

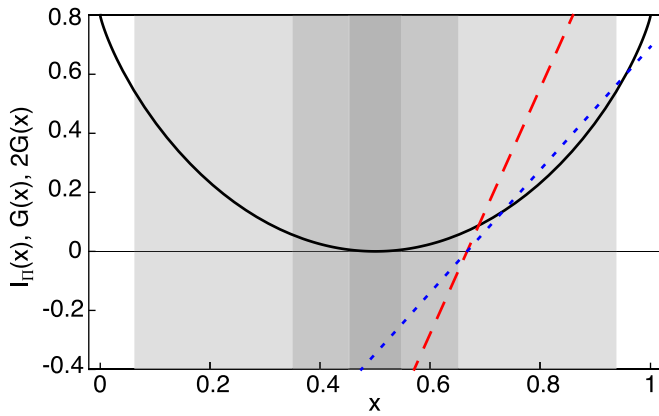
$$b(\lambda, t) \approx \frac{\sup_{x \in [x_-, x_+]} [2G(x) - I_{\Pi}(x)]}{\sup_{x \in [x_-, x_+]} [G(x) - I_{\Pi}(x)]}, \quad [13]$$

where the dependence on  $t$  is through  $x_+$  and  $x_-$ . The zero of the rate function,  $x_{\min} = 1/2$ , corresponds to the most probable value of the product in Eq. 1. Because  $x_{\pm} \approx 1/2 \pm \sqrt{((1-\lambda)/2\lambda)((\log R)/t)}$ , for fixed  $R$  the suprema in Eq. 13 are computed over an increasingly narrower set around  $x_{\min}$  [with  $I_{\Pi}(x_{\min}) = 0$ ] as  $t$  increases (Fig. 5). Thus, for any finite number of realizations

**Table 2.** Statistics of estimated sample exponents in the interspecific generalized TL on carabid beetles abundances

$j, k$	$k/j$	$b_{jk}$	2.5% percentile	97.5% percentile
1, 2	2	2.005	1.984	2.025
1, 3	3	3.005	2.961	3.042
1, 4	4	3.994	3.936	4.057

The column  $k/j$  gives the asymptotic model prediction for the exponent  $b_{jk}$ . The point estimate  $b_{jk}$  is computed as the average  $b_{jk}$  across years and site type, not by pooling all of the data from different years and site types to calculate means and variances. The confidence intervals are obtained via bootstrapping with  $10^6$  bootstrap samples from the set of  $b_{jk}$ .



**Fig. 5.** Plot of  $I_{\Pi}(x)$  (black curve),  $G(x)$  (dotted blue line), and  $2G(x)$  (dashed red line). Marked in gray are the regions  $[x_-, x_+]$  at times  $t = 10, 100$ , and  $1,000$  (from light to dark gray) for fixed  $R = 100$ . These gray regions are the intervals over which the supremum in Eq. 13 is computed. In this example,  $r = 2, s = 1/4, \lambda = 0.55$ . The quantities  $x_+$  and  $x_-$  are computed by solving numerically Eq. 12.

$R$ , the sample exponent will approximate  $\lim_{t \rightarrow \infty} b(\lambda, t) \simeq 2$  after a time  $t^*$  that increases only logarithmically with  $R$  (Eq. 12 and Fig. 2), for any choice of  $\lambda, r$ , and  $s$ . For example, with  $\lambda = 0.5$ , when  $t = 100$ , to access the extreme event  $x_+ = 0.9$  (and  $x_- = 0.1$ ) one needs about  $R \simeq 10^{13}$  replicates of the process. More precisely, the sample exponent is close to the population exponent if the arguments of the two suprema in Eq. 9 are included in  $[x_-, x_+]$ . Using  $x_{\pm} \simeq 1/2 \pm \sqrt{((1-\lambda)/2\lambda)((\log R)/t)}$ , if the largest argument of the suprema in Eq. 9 is  $\bar{x} > 1/2$ , the sample exponent approximates accurately the population exponent for  $t < ((1-\lambda)/\lambda)(\bar{x} - 1/2)^{-2} \log R$  and tends to  $k/j$  for  $t > ((1-\lambda)/\lambda)(\bar{x} - 1/2)^{-2} \log R$  (Figs. 1 and 2). If  $r$  and  $s$  are such that the population TL exponent  $b$  displays a discontinuity at  $\lambda = \lambda_c$  [in which case  $\lambda_c = (1-r-s+rs)/(-r-s+2rs)$ ] (SI Methods), then the above results give the minimum number of replicates required to observe such discontinuity also in the sample TL exponent.

Analogous considerations hold for the asymptotic sample exponent describing the scaling of the sample moments  $\mathbb{E}[N(t)^k]$  with  $\mathbb{E}[N(t)^j]$ , which can be approximated as

$$b_{jk}(\lambda, t) \simeq \frac{\sup_{x \in [x_-, x_+]} [kG(x) - I_{\Pi}(x)]}{\sup_{x \in [x_-, x_+]} [jG(x) - I_{\Pi}(x)]}, \quad [14]$$

which is the analog of Eq. 13 for any pair of sample moments. Fig. S1 C and D shows that simulation results and theoretical predictions for  $b_{jk}$  show excellent agreement in the two regimes  $t \gg \log R$  and  $1 \ll t \ll \log R$ .

A standard saddle-point calculation suggests that the limiting growth rate of the variance is equal to the limiting growth rate of the second moment also for ergodic transition matrices, apart from peculiar cases (see ref. 25 for a discussion of a counterexample). The same argument suggests that the limiting growth rate of the  $k$ th cumulant equals that of the  $k$ th moment ( $t^{-1} \log \mathbb{E}[N(t)^k]$ ) for large  $t$ . The suggested equivalence between the scaling exponents of cumulants and moments for ergodic II would allow extending the result on the sample TL ( $b = 2$ ) and generalized TL ( $b = k/j$ ) to the scaling of cumulants in  $m$ -step Markov chains, whose transition matrix is ergodic but not twofold irreducible. However, pathological counterexamples may exist.

Eq. 13 gives the estimated sample exponent of TL asymptotically, ignoring the constant term in the scaling of the variance  $V$  vs. the mean  $M$  as  $\log V = b \log M + \log a$ . For small  $t$ ,  $\log a$  can be of the same order of magnitude as  $\log V$ . Fig. 2 shows the crossover of the sample exponent (for fixed  $R, \lambda, r$ , and  $s$ ) from the population exponent  $b = b(\lambda)$  as in Eq. 9 (observed when  $t \ll \log R$ ) to  $b \simeq 2$  (when  $t \gg \log R$ ), where the sample exponent is calculated as the slope of the curve  $\log \mathbb{E}[N(t)^2]$  vs.  $\log \mathbb{E}[N(t)]$  at time  $t$  (thus not neglecting the constant term  $\log a$ ). The sample moments are computed as

$$t^{-1} \log \mathbb{E}[N(t)^k] \simeq \sup_{x \in [x_-, x_+]} [kG(x) - I_{\Pi}(x)] \quad [15]$$

(compare Eq. 8) in Fig. 2A and as the sample moments in simulations in Fig. 2B.

See SI Methods for further details and generalizations.

**Empirical Analysis.** We used the BRF dataset to show that the generalized TL holds with sample exponent  $b_{jk} = k/j$ . We computed the moment ratios  $\langle [N(t)/N_0]^k \rangle$ , where the symbol  $\langle \cdot \rangle$  identifies the sample mean across the six plots of BRF and  $N_0$  is the number of trees at the start of the census in 1931. Following ref. 5, we tested whether the moments of the spatial density ratio  $N(t)/N_0$  in the five most recent censuses satisfied TL and the generalized TL with  $b_{jk} = k/j$ . Table 1 reports the slopes of the least-squares linear regressions of  $\langle [N(t)/N_0]^k \rangle$  vs.  $\langle [N(t)/N_0]^j \rangle$ , which are all compatible with the model prediction  $b_{jk} = k/j$ . The BRF dataset thus provides an empirical example where the multiplicative model satisfactorily describes the underlying dynamics and the generalized TL holds asymptotically as the model predicts.

The intraspecific form of TL and the generalized scaling relationship between higher moments (Eq. 2) were tested using abundance data from 26 species of carabid beetles. We have limited the analysis of the intraspecific TL to the set of species that were present in all sites in each given year. We have followed the researchers who collected the carabid beetles abundance data (30) in excluding species with year samples with zero individuals in at least one of the sites from the statistical analysis. The authors of ref. 30 declared that they were unable to differentiate sites where a species was not present from sites where the density of such species was so low that no catches were realized. For each species, we selected data from a minimum of three to a maximum of six sites (all either woodland or heath) (30) and from a minimum of 4 to a maximum of 6 consecutive years. The precise number of sites and years varied for each species, depending on the number of sites and years in which at least one individual of such species was found in each site. The moments of species abundance were calculated separately for each species and for each available year. Linear regressions of  $\log(N^k(t))$  vs.  $\log(N^j(t))$  for  $j = 1, 2, \dots, Y$  [where  $Y$  is the total number of available years for the selected species and  $\langle N^k(t) \rangle$  is the  $k$ th spatial sample moment in year  $t$ ] gave the estimate of the sample exponent  $b_{jk}$  for the selected species (Fig. 3).

The interspecific form of TL and the generalized scaling relationship for statistical moments (Eq. 2) were investigated following ref. 16, using the carabid beetles dataset, computing spatial sample moments across similar sites. Data from sites labeled B, C, X, and AE in ref. 30, collected between 1961 and 1966, were used to calculate spatial moments across woodland sites. Data from sites labeled AT, N, Z, and AG in ref. 30, collected between 1963 and 1966, were used to calculate spatial moments across heath sites. As for the intraspecific TL analysis, we have limited the analysis of the interspecific TL to the set of species that were present in all sites in each given year. Spatial moments of carabid beetles abundance were computed for each species individually and separately for each year and site type (woodland or heath). For each year, we calculated the least-squares slope of  $\log(N^k)$  vs.  $\log(N)$  across all species at a given year and site type. Tables S1 and S2 show the summary statistics for all years and site types. Fig. 4A and Fig. S5 M and N show the scaling of the  $k$ th sample moment  $\langle N^k \rangle$  with  $\langle N \rangle$  when data for all years and site types are plotted together; each data point in Fig. 4 and in Fig. S5 refers to the spatial moments of a single species in one year and site type. Fig. S5 A–L shows the scaling of the  $k$ th sample moment  $\langle N^k \rangle$  with  $\langle N \rangle$  for each year and site type separately. The least-squares exponents  $b_{jk}$  computed in the linear regression of  $\log(N^k)$  vs.  $\log(N^j)$  are compatible with the asymptotic model prediction  $b_{jk} = k/j$  (Tables S1 and S2), as are the mean exponents  $b_{jk}$  (Table 2).

See SI Methods and Tables S3 and S4 for further details.

**ACKNOWLEDGMENTS.** We thank Dr. Hugo Touchette for discussions and Dr. Markus Fischer for discussions and a careful reading of the manuscript. A.R. and A.G. acknowledge the support provided by the discretionary funds of Eawag: Swiss Federal Institute of Aquatic Science and Technology and by the Swiss National Science Foundation Project 200021\_157174. M.F. has been partially supported by Grantová agentura České republiky Grant P201/12/2613. J.E.C. acknowledges the support of US National Science Foundation Grant DMS-1225529 and the assistance of Priscilla K. Rogerson.

1. Taylor LR (1961) Aggregation, variance and the mean. *Nature* 189(4766):732–735.
2. Marquet PA, et al. (2005) Scaling and power-laws in ecological systems. *J Exp Biol* 208(Pt 9):1749–1769.
3. Fronczak A, Fronczak P (2010) Origins of Taylor's power law for fluctuation scaling in complex systems. *Phys Rev E Stat Nonlin Soft Matter Phys* 81(6 Pt 2):066112.

4. Ramsayer J, Fellous S, Cohen JE, Hochberg ME (2012) Taylor's Law holds in experimental bacterial populations but competition does not influence the slope. *Biol Lett* 8(2):316–319.
5. Cohen JE, Xu M, Schuster WSF (2013) Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. *Proc Biol Sci* 280(1757):20122955.



# Supporting Information

Giometto et al. 10.1073/pnas.1505882112

## SI Methods

In *SI Methods* and *SI Text*, the symbol  $b$  refers to the conventional TL population exponent, the symbol  $b_R$  refers to the conventional TL sample exponent, and the generalized TL exponents are indicated with the symbol  $b_{jk}$  (the distinction between sample and population exponents will be clear from the context). Both sample and population exponents were indicated as  $b$  (or  $b_{jk}$ ) in the main text to simplify the notation. The calculations reported in *Methods* (main text) identify the logarithmic dependence of  $x_+$  on the number of realizations  $R$ , but rely on a number of approximations: the definition of  $x_+$  (which, in a given realization, is a random variable), the computation of Laplace integrals (Eq. 11), and the expansion of the rate function around  $x_{\min}$  (Eq. 12). Such calculations can be made more rigorous if we consider the independent identically distributed random variables  $X^i(t) = L_i^i(r)$ ; that is,  $X^i(t)$  is the frequency of occurrence of the first state up to time  $t$  in the  $i$ th realization of the Markov chain ( $i = 1, \dots, R$ ). We now define  $x^+ = \max\{X^1(t), \dots, X^R(t)\}$  and observe that

$$\frac{1}{t} \log \mathbb{P}[X^1(t) > x] \leq \frac{1}{t} \log \mathbb{P}[x^+ > x] \leq \frac{1}{t} \log \mathbb{P}[X^1(t) > x] + \frac{1}{t} \log(R). \quad \text{[S1]}$$

Note that all logarithms here and in the main text are to the base  $e$ . For fixed  $R$  [or, more generally,  $\log R = o(t)$ ] and  $x > 1/2$ , taking the limit ( $\lim_{t \rightarrow \infty}$ ) in Eq. S1 and knowing that  $L^i(r)$  satisfies a LDP, one has

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P}[x^+ > x] = \sup_{y \in (x, 1]} -I_{\Pi}(y) = -I_{\Pi}(x). \quad \text{[S2]}$$

Because  $0 < I_{\Pi}(x) \leq \infty$ , Eq. S2 implies that  $\lim_{t \rightarrow \infty} \mathbb{P}(x^+ > x) = 0$  for any  $x > 1/2$ . An analogous calculation for  $x^- = \min\{X^1(t), \dots, X^R(t)\}$  shows that  $\lim_{t \rightarrow \infty} \mathbb{P}(x^- < x) = 0$  for any  $x < 1/2$ . In this context, we can approximate the sample exponent at time  $t$  with an analog of Eq. 13:

$$b_R(\lambda, t) \simeq \frac{\sup_{x \in [x^-, x^+]} [2G(x) - I_{\Pi}(x)]}{\sup_{x \in [x^-, x^+]} [G(x) - I_{\Pi}(x)]}. \quad \text{[S3]}$$

In the narrow interval  $[x^-, x^+]$  centered around  $x_{\min}$ ,  $I_{\Pi}(x) \simeq 0$  and as a consequence  $b_R(\lambda, t) \simeq 2$  (Fig. 5). More precisely,  $|b_R(\lambda, t) - 2|$  goes to 0 in probability as  $t$  tends to infinity. In fact, for every  $\epsilon > 0$ , we have

$$\mathbb{P}[|b_R(\lambda, t) - 2| > \epsilon] \leq \mathbb{P}\left[x^+ > \frac{1}{2} + \eta(\epsilon)\right] + \mathbb{P}\left[x^- < \frac{1}{2} - \eta(\epsilon)\right], \quad \text{[S4]}$$

where  $\eta(\epsilon)$  is a function that goes to zero for  $\epsilon \rightarrow 0$ . Because of Eqs. S1 and S2, it follows that

$$\lim_{t \rightarrow \infty} \mathbb{P}[|b_R(t) - 2| > \epsilon] = 0. \quad \text{[S5]}$$

Analogous considerations hold for the generalized TL describing the scaling of any pair of moments.

We now look at some generalizations of the stochastic multiplicative process considered above. The sample exponent in a finite set of  $R$  independent realizations of the process is  $b \simeq 2$

also for nonsymmetric transition matrices  $\Pi$ . In the asymmetric case, the transition matrix is

$$\Pi = \begin{pmatrix} 1 - \lambda & \lambda \\ \mu & 1 - \mu \end{pmatrix}, \quad \text{[S6]}$$

with  $0 < \lambda, \mu < 1$ . The rate function  $I_{\Pi}(x)$  is convex and attains its minimum at  $x_{\min} = \pi(1) = \mu/(\lambda + \mu)$ , where  $\pi = (\pi(1), \pi(2) = \lambda/(\lambda + \mu))$  is the invariant measure for  $\Pi$  and  $I_{\Pi}(x_{\min}) = 0$ . Only the value of the rate function at  $x_{\min}$  and not the value of  $x_{\min}$  is relevant for our argument. Due to asymmetries of  $I_{\Pi}$ , “left” (i.e.,  $x < x_-$ ) rare events could be easier to see than “right” (i.e.,  $x > x_+$ ) rare events or vice versa. In all cases, an exponentially large in  $t$  number of replicates is needed to sample the tails with the correct weights. In this context, Eqs. 9, 10, 13, and 14 and Eq. S3 are still valid and give, respectively, the asymptotic population and sample exponents.

The previous considerations can also be extended to multiplicative processes  $N(t)$  in more general Markovian environments with  $w$  states and state space  $\chi = \{r_1, \dots, r_w\}$ , where all  $r_i$  are strictly positive and at least two  $r_i$  are different. We label the state space  $\chi = \{1 \leftrightarrow r_1, \dots, w \leftrightarrow r_w\}$ . Let the transition matrix  $\Pi$  be twofold irreducible (i.e.,  $\Pi$  irreducible and  $\Pi \Pi^T$  irreducible, where  $\Pi^T$  is the transpose of  $\Pi$ ). The rate function in Eq. 4 reads (theorem IV.7 and section IV.3 of ref. 1 or theorem 3.1.6 of ref. 2)

$$I_{\Pi}(\mu) = \sup_{u > 0} \left[ \sum_{v=1}^w \mu_v \log \frac{u_v}{(\Pi u)_v} \right], \quad \text{[S7]}$$

where  $u$  is a strictly positive vector in  $\mathbb{R}^w$ . Here,  $\sum_{v=1}^w \mu_v = 1$ , and  $\mu_v$  represents the proportion of  $v$  after  $t$  steps (for large  $t$ ). The rate function is convex and  $I_{\Pi}(\mu_{\min}) = 0$ , with  $\mu_{\min}$  the most probable state for large  $t$  (theorems 3.1.2 and 3.1.6 of ref. 2 and section 4.3 of ref. 3). Eq. 9, with  $x$  in the standard  $w - 1$  simplex in  $\mathbb{R}^w$  and  $G(x) = \sum_{i=1}^w x_i \log r_i$ , gives the population scaling exponent of  $\mathbb{E}[N(t)^2]$  with  $\mathbb{E}[N(t)]$ . The twofold irreducibility of  $\Pi$  plus the condition that  $r_i \neq r_j$  for some  $i \neq j$  is the sharpest sufficient assumption that is presently known (4) to guarantee that the limiting growth rate of the second moment equals the limiting growth rate of the variance; thus, Eq. 9, with  $x$  in the standard  $w - 1$  simplex in  $\mathbb{R}^w$  and  $G(x) = \sum_{i=1}^w x_i \log r_i$ , gives the population scaling exponent of  $\text{Var}[N(t)]$  with  $\mathbb{E}[N(t)]$ . Analogously, Eq. 10, with  $x$  in the standard  $w - 1$  simplex in  $\mathbb{R}^w$  and  $G(x) = \sum_{i=1}^w x_i \log r_i$ , gives the population scaling exponent of  $\mathbb{E}[N(t)^k]$  with  $\mathbb{E}[N(t)^j]$ . As far as the scaling of moments is of interest, the ergodicity (i.e., irreducibility and aperiodicity) of  $\Pi$  (as opposed to the twofold irreducibility) and  $G(x)$  not identically equal to zero (which happens only if  $r_i = 1 \forall i$ ) are sufficient to compute the scaling exponents via Eqs. 9 and 10, modified as stated above. This is true because the ergodicity of  $\Pi$  ensures that the empirical measure  $L_t$  satisfies a LDP (theorems 3.1.2 and 3.1.6 of ref. 2). Therefore, one can apply Varadhan’s lemma (theorem III.13 of ref. 1) to compute the limiting growth rate of the moments of  $N(t)$  via Eq. 8, with  $x$  in the standard  $w - 1$  simplex in  $\mathbb{R}^w$  and  $G(x) = \sum_{i=1}^w x_i \log r_i$ . The computation of the sample exponents  $b_R$  and  $b_{jk}$  is similar to that in the two-state case and the sample exponents approximate  $b_R = 2$  and  $b_{jk} = k/j$  asymptotically in time; the proof is as follows. We consider the independent identically distributed random variables  $Y^i(t) = |L_t^i - \mu_{\min}|$ , where  $L_t^i = (L_t^i(r_1), \dots, L_t^i(r_w))$  and the superscript  $i$  indicates the  $i$ th

independent realization of the chain ( $i = 1, \dots, R$ ). We now define  $y^+ = \max\{Y^1(t), \dots, Y^R(t)\}$  and observe that, for every  $\epsilon > 0$ ,

$$\mathbb{P}(y^+ > \epsilon) \leq R \mathbb{P}(Y^1(t) > \epsilon). \quad [\text{S8}]$$

For fixed  $R$  and  $\epsilon$ , taking the limit ( $\lim_{t \rightarrow \infty}$ ) in Eq. S8 and knowing that  $L_t^1$  satisfies a LDP (in particular,  $\lim_{t \rightarrow \infty} \mathbb{P}(Y^1(t) > \epsilon) = 0$ ), one has

$$\lim_{t \rightarrow \infty} \mathbb{P}(y^+ > \epsilon) = 0. \quad [\text{S9}]$$

In this context, we can approximate the sample exponent with

$$b_R(\lambda, t) \simeq \frac{\sup_{|\mu - \mu_{\min}| < y^+} [2G(\mu) - I_{\Pi}(\mu)]}{\sup_{|\mu - \mu_{\min}| < y^+} [G(\mu) - I_{\Pi}(\mu)]}. \quad [\text{S10}]$$

In the narrow region  $|\mu - \mu_{\min}| < y^+$  centered around  $\mu_{\min}$ ,  $I_{\Pi}(\mu) \simeq 0$  and as a consequence  $b_R(\lambda, t) \simeq 2$ . More precisely,  $|b_R(\lambda, t) - 2|$  goes to 0 in probability as  $t$  tends to infinity. In fact, for every  $\delta > 0$ , we have

$$\mathbb{P}(|b_R(\lambda, t) - 2| > \delta) \leq \mathbb{P}(y^+ > \eta(\delta)), \quad [\text{S11}]$$

where  $\eta(\delta)$  is a function that goes to zero for  $\delta \rightarrow 0$ . Because of Eq. S9, it follows that

$$\lim_{t \rightarrow \infty} \mathbb{P}(|b_R(t) - 2| > \delta) = 0. \quad [\text{S12}]$$

Analogous considerations hold for the generalized TL describing the scaling of any pair of moments. A standard saddle-point calculation suggests that the limiting growth rate of the variance is equal to the limiting growth rate of the second moment also for ergodic transition matrices, apart from peculiar cases (see ref. 4 for a discussion of a counterexample). The same argument suggests that the limiting growth rate of the  $k$ th cumulant equals that of the  $k$ th moment ( $t^{-1} \log \mathbb{E}[N(t)^k]$ ) for large  $t$ . The suggested equivalence between the scaling exponents of cumulants and moments for ergodic  $\Pi$  would allow extending the result on the sample TL ( $b = 2$ ) and generalized TL ( $b = k/j$ ) to the scaling of cumulants in  $m$ -step Markov chains, whose transition matrix is ergodic but not twofold irreducible. However, pathological counterexamples may exist.

**Analysis of the discontinuity in  $b$  as a function of  $r$  and  $s$ .** A discontinuity in the population TL exponent  $b$  (Fig. 1 and Eq. 9) is present when the limiting growth rate of the mean abundance is zero; i.e.,  $\lim_{t \rightarrow \infty} (1/t) \log \mathbb{E}[N(t)] = 0$ . Let us consider Fig. 5 and fix  $r$  and  $s$  with  $r \neq s$ . The value of  $\lambda$  shapes the form of  $I_{\Pi}(x)$  (black curve in Fig. 5); in particular, the second derivative can be easily calculated from Eq. 5 and shown to increase for larger  $\lambda$ . A discontinuity may eventually appear for the value  $\lambda = \lambda_c$  such that the curve  $I_{\Pi}(x)$  and the line  $G(x)$  (blue line in Fig. 5) are tangent. In other words,  $\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)] = \sup_{x \in [0,1]} [G(x) - I_{\Pi}(x)] = 0$  for  $\lambda = \lambda_c$  such that

$$\left. \begin{aligned} & \log \left\{ \frac{1}{2} [(1 - \lambda_c)(r + s) \right. \\ & \left. + \sqrt{4(2\lambda_c - 1)rs + (\lambda_c - 1)^2(r + s)^2}] \right\} = 0, \end{aligned} \quad [\text{S13}]$$

with constraints  $r, s > 0$  and  $0 < \lambda_c < 1$ .  $\lambda_c$  exists only for certain values of  $r$  and  $s$ , and thus a discontinuity in the population TL exponent  $b$  is not always possible. Solving Eq. S13 with respect to  $\lambda_c$  gives  $\lambda_c = (1 - r - s + rs)/(-r - s + 2rs)$ ; thus, for any given  $s$ ,  $\lambda_c = 0$  for  $r = 1$  and  $\lambda_c = 1$  for  $r = 1/s$ . For fixed  $s \neq 1$  one has

$d\lambda_c/dr > 0$  (except for  $r = s$  where  $d\lambda_c/dr|_{r=s} = 0$ ); thus,  $\lambda_c$  exists for  $0 < r \leq 1/s$  and  $r \geq 1$  if  $s > 1$  and for  $1 \leq r \leq 1/s$  if  $s < 1$  (Fig. S3). Fig. S4 schematically illustrates the behavior of  $b(\lambda)$  for different pairs  $\{r, s\}$  of multiplicative factors. Discontinuities analogous to that of  $b(\lambda)$  appear for certain values of  $r, s$ , and  $\lambda$  in the population exponents  $b_{jk}$  (Eq. 10), when  $\lim_{t \rightarrow \infty} t^{-1} \log \mathbb{E}[N(t)^j] = \sup_{x \in [0,1]} [jG(x) - I_{\Pi}(x)] = 0$ .

#### Transient Agreement Between Sample and Population Exponents.

Sample and population exponents may display transient agreement in the regime  $1 \ll t \ll \log R$  (e.g., black solid lines and black solid circles in Fig. 1 or the red solid curve in Fig. 2B), if the number of replicates  $R$  is not too small. However, population exponents were proved to obey Taylor's law only asymptotically in time (compare Eqs. 8 and 9 here and theorem 1 of ref. 4). To understand the cause of such an agreement, one needs to consider two different asymptotic regimes:

- i) The first regime, discussed in the main text, is the asymptotic regime  $t \gg \log R$ . In such a regime, rare events are not accessible and the sample exponents are not representative of the population exponents. We call this regime the asymptotic sample regime, which we have proved to result in a constant sample TL exponent  $b = 2$ .
- ii) The second asymptotic regime concerns the fact that population exponents were proved to obey Taylor's law only asymptotically in time. We refer to the second regime as the asymptotic population regime. Population exponents take into account all possible realizations of the process, including rare events. If  $R$  (fixed) is not too small, increasing  $t$  from  $t = 1$ , the asymptotic population regime may occur earlier in time than the asymptotic sample regime. In this case one can observe rare events with proper statistics and Eqs. 9 and 10 give a good prediction of both sample and population TL exponents, as long as  $t \ll \log R$ .

The red solid curve in Fig. 2B exemplifies the above discussion: Initially, the sample TL exponent is different from the theoretical asymptotic prediction for the population TL exponent (Fig. 2B, dashed upper horizontal line). For  $3 \leq t \leq 12$ , the asymptotic prediction for the population TL exponent (Fig. 2B, dashed upper horizontal line) gives a good prediction for the sample TL exponent. Finally, for large  $t$ , the sample TL exponent approximates the asymptotic prediction  $b \simeq 2$ .

**Compatibility of Eq. 9 here and Equation 8 in Ref. 4.** We show here that Eq. 9 coincides with equation 8 in ref. 4, under the assumption (stronger than in ref. 4) that the transition matrix  $\Pi$  is positive and  $r \neq s$ . The rate function Eq. 4 can be written as (section 4.3 of ref. 3 or theorem 3.1.7 of ref. 2)  $I_{\Pi}(x) = \sup_q \{qx - \log \zeta(\Pi_q)\}$ , where  $\Pi_q$  is the matrix with elements  $\Pi_q(i, j) = \Pi(i, j) \exp(q\delta_{j,1})$ , and  $\zeta(\cdot)$  indicates the spectral radius (i.e., the Perron-Frobenius eigenvalue).  $\zeta(\Pi_q)$  is unique and analytic in  $q$ ; thus,  $\xi(q) \equiv \log \zeta(\Pi_q)$  is differentiable and the rate function can be expressed as  $I_{\Pi}(x) = q(x)x - \xi(q(x))$ , where  $q(x)$  is the unique solution of  $\xi'(q) = x$ . Eq. 8 for the  $k$ th moment of  $N(t)$  then reads  $\lim_{t \rightarrow \infty} (1/t) \log \mathbb{E}[N(t)^k] = \sup_{x \in [0,1]} [kG(x) - q(x)x + \xi(q(x))]$ . The argument of the supremum is maximum at  $x^*$  such that  $k \log(r/s) - q(x^*) = 0$ ; that is,  $x^* = \xi'(k \log(r/s))$ . Thus, evaluating the supremum, one has  $\lim_{t \rightarrow \infty} (1/t) \log \mathbb{E}[N(t)^k] = k \log s + \xi(k \log(r/s)) = \log [s^k \zeta(\Pi_k \log(r/s))] = \log \zeta(\Pi \text{diag}(r, s)^k)$ , which coincides with equations 13 and 14 of ref. 4 [equations 13 and 14 in ref. 4 are expressed in terms of the column-stochastic matrix  $\Pi^T$  that corresponds to the row-stochastic matrix  $\Pi$ ; because  $\zeta(\text{diag}(r, s)^k \Pi^T) = \zeta(\Pi \text{diag}(r, s)^k)$ , the equations coincide]. The compatibility of Eq. 9 here with equation 8 in ref. 4 follows directly.

**Software and Numerical Analysis.** Simulation of the multiplicative process in Eq. 1 in software with finite precision is subject to

numerical underflow and overflow. This may result in errors in the estimation of exponentially growing or declining abundances after very few steps, if simulations are not performed carefully. For simulations performed in this study we used a symbolic software that allows infinite precision calculations and thus simulates correctly the multiplicative process in Eq. 1 and computes exactly the moments at every time  $t$ . Therefore, all numerical calculations in this study are free of underflow and overflow issues.

**Generalized TL for Tree Abundance in the Black Rock Forest (USA).** We tested the predictions of the multiplicative growth model by using a dataset of tree abundance from six long-term plots in the Black Rock Forest (BRF), Cornwall, New York. We computed the moment ratios  $\langle [N(t)/N_0]^k \rangle$ , where the symbol  $\langle \cdot \rangle$  identifies the sample mean across the six plots of BRF and  $N_0$  is the number of trees at the start of the census in 1931. Following ref. 5, we tested whether the moments of the spatial density ratio  $N(t)/N_0$  in the five most recent censuses satisfied TL and the generalized TL with  $b_{jk} = k/j$ . Table 1 reports the slopes of the least-squares linear regressions of  $\langle [N(t)/N_0]^k \rangle$  vs.  $\langle [N(t)/N_0]^j \rangle$ , which are all compatible with the model prediction  $b_{jk} = k/j$ . The BRF dataset thus provides an empirical example where the multiplicative model satisfactorily describes the underlying dynamics and the generalized TL holds asymptotically as the model predicts.

**Generalized TL for Carabid Beetles Abundance.** Here, we test the multiplicative growth model hypotheses on the carabid beetles dataset. The carabid beetles dataset consists of abundance data of carabid beetles ranging from a minimum of three to a maximum of six sites and from a minimum of 4 to a maximum of 6 consecutive years, depending on the species. We computed the multiplicative factors  $A(p, t) = N(p, t)/N(p, t-1)$  separately for each species, site  $p$ , and pair of consecutive years. We tested some of the assumptions of the multiplicative growth model, namely the independence and identical distribution of multiplicative factors over sites and over time. Each test was performed separately for each species. The tests used rely on assumptions, such as normality of data, which were tested before performing the hypothesis testing. We excluded from such tests the species for which the test assumptions were not met. Tables S3 and S4 report the percentage of species for which a  $P$  value smaller than 0.05 was returned, when testing for the identical distribution of multiplicative factors over sites and time, respectively. The number of species used in each test, that is, the number of species that met the test assumptions, is reported in the third column of Tables S3 and S4. The first four tests in Tables S3 and S4 test for identical mean and the last four tests test for identical variance. The high percentages of rejection of the null hypotheses of equal mean and equal variance of multiplicative factors over sites and time in the carabid beetles dataset suggest that the carabid beetles population dynamics do not conform to the Markovian multiplicative growth model. Nevertheless, the predictions of the analysis regarding the higher-order sample exponents of the generalized TL were substantially confirmed. That the generalized TL pattern holds in the carabid beetles dataset, despite the disagreement with the assumptions of the Markovian multiplicative model, suggests that the results of our theoretical investigation might hold far beyond the population growth model considered in the main text.

## SI Text

**Comparison with Other Demographic Models.** The multiplicative growth model is one of numerous demographic models that predict TL. The exponent  $b = 2$  for the scaling of the variance vs. the mean is typical of deterministic dynamics. For example, an exponential model of clonal growth (6), where clones grow exponentially with different growth rates (variability enters here only through the different growth rates and initial densities), and

the above symmetric model for  $\lambda = 0$  or  $\lambda = 1$  both predict TL with exponent  $b = 2$ . Although found in deterministic models, the exponent  $b = 2$  is also observed in stochastic models such as the continuous-time birth–death process and the Galton–Watson branching process (4). Such models yield population exponents  $b = 2$  and  $b = 1$ , respectively, for asymptotically growing and decaying populations (4).

The theoretical investigation of multiplicative population processes showed that the generalized TL sample exponents  $b_{jk}$  satisfy  $b_{jk} \simeq k/j$  asymptotically for large  $t$  for a broad ensemble of transition matrices  $\Pi$  and sets of positive multiplicative factors. Additionally, our large-deviation approach and our small-sample argument suggest that the entropic term in Eq. 13 dominates over the other terms that contain the specifications of the demographic process. Thus, the result might be more general than the class of multiplicative population growth models. We show here that  $b_{jk} = k/j$  holds for the population exponents of other population growth processes, such as the birth–death process in the case of expanding populations.

The moments of the birth–death process with constant birth rate  $\lambda$  and constant death rate  $\mu$  can be computed via the associated moment-generating function  $M$ , which is equal to (7)

$$M(\theta, t) = \left( \frac{\mu v(\theta, t) - 1}{\lambda v(\theta, t) - 1} \right)^{N_0}, \quad [\text{S14}]$$

where  $v(\theta, t) = (e^\theta - 1)e^{(\lambda - \mu)t} / (\lambda e^\theta - \mu)$  and  $N_0$  is the initial population size. The  $k$ th moment of population size can be computed as  $\langle N^k \rangle = (\partial^k M(\theta, t) / \partial \theta^k) |_{\theta=0}$ . Here, we assume  $N_0 = 1$  (but the result holds for any  $N_0$ ) and an expanding population; i.e.,  $\lambda - \mu > 0$ . Because  $v(0, t) = 0$ ,  $(\partial v / \partial \theta)(\theta, t) = (\lambda - \mu)e^{(\lambda - \mu)t} (e^\theta / (-e^\theta \lambda + \mu)^2) \propto e^{(\lambda - \mu)t}$ , and  $(\partial^k v / \partial \theta^k)(\theta, t) \propto e^{(\lambda - \mu)t}$ , the leading term in the partial derivatives of  $M(\theta, t)$  with respect to  $\theta$ , evaluated in  $\theta = 0$ , can be written as

$$\begin{aligned} \frac{\partial^k M}{\partial \theta^k}(\theta, t) \Big|_{\theta=0} &= (-1)^{k+1} (\lambda - \mu) \lambda^{k-1} \frac{(\partial v / \partial \theta)^k}{(-1 + \lambda v)^{k+1}} \Big|_{\theta=0} \\ &+ o \left[ \left( \frac{\partial v}{\partial \theta} \right)^k \Big|_{\theta=0} \right] \\ &= (\lambda - \mu)^{1-k} \lambda^{k-1} e^{k(\lambda - \mu)t} + o \left[ e^{k(\lambda - \mu)t} \right], \end{aligned} \quad [\text{S15}]$$

where the lowercase- $o$  notation indicates that the remaining terms are negligible in the limit  $t \rightarrow \infty$ . Derivation of the equation for  $(\partial^k M / \partial \theta^k)(\theta, t)$  (first line of Eq. S15) shows that the leading term in  $(\partial^{k+1} M / \partial \theta^{k+1})(\theta, t) |_{\theta=0}$  is equal to  $(\lambda - \mu)^k \lambda^k e^{(k+1)(\lambda - \mu)t} + o[e^{(k+1)(\lambda - \mu)t}]$ , which coincides with replacing  $k$  by  $k + 1$  in Eq. S15. Eq. S15 can be obtained by considering that, because  $\partial^k v / \partial \theta^k \propto e^{(\lambda - \mu)t}$  and  $v(0, t) = 0$ , the leading term in  $(\partial M / \partial \theta)(\theta, t) = (\lambda - \mu)((\partial v / \partial \theta) / (-1 + \lambda v)^2)$  evaluated at  $\theta = 0$  is the second term in the quotient rule  $(f/g)' = (f'g - fg')/g^2$ , that is, the term that raises the exponent of  $\partial v / \partial \theta$  by 1 unit. For subsequent derivatives, the quotient rule is applied to the leading term. All other terms in  $(\partial^k M / \partial \theta^k)(\theta, t) |_{\theta=0}$  contain products of partial derivatives; for example,

$$\frac{\partial^2 M}{\partial \theta^2} \Big|_{\theta=0} = (\lambda - \mu) \left( 2\lambda \frac{\partial v}{\partial \theta} \Big|_{\theta=0} + \frac{\partial^2 v}{\partial \theta^2} \Big|_{\theta=0} \right), \quad [\text{S16}]$$

$$\frac{\partial^3 M}{\partial \theta^3} \Big|_{\theta=0} = (\lambda - \mu) \left( 6\lambda^2 \frac{\partial v}{\partial \theta} \Big|_{\theta=0} + 6\lambda \frac{\partial v}{\partial \theta} \frac{\partial^2 v}{\partial \theta^2} \Big|_{\theta=0} + \frac{\partial^3 v}{\partial \theta^3} \Big|_{\theta=0} \right), \quad [\text{S17}]$$

i.e.,  $\prod_{j=1}^k (\partial^j v / \partial \theta^j)^{q_j}$ , with  $\sum_{j=1}^k q_j < k$  (with  $q_j \in \mathbb{N}$ ), and are thus negligible in the limit  $t \rightarrow \infty$ . From Eq. S15 it follows that

$\lim_{t \rightarrow \infty} (1/t) \log \langle N^k \rangle = k(\lambda - \mu)$ ; thus, the generalized TL holds with  $b_{jk} = k/j$ .

The asymptotic behavior of exponents  $\lim_{t \rightarrow \infty} (1/t) \log \langle N^k \rangle = k(\lambda - \mu)$  can also be computed via the continuous approximation of the birth–death process. Although such calculations do not provide further understanding of the birth–death process (we have already calculated the limiting behavior of  $\langle N^k \rangle$  for large  $t$ ), the fact that the continuous approximation of the birth–death process coincides with that of the Galton–Watson branching process (8–10) suggests an even broader validity for the generalized TL result  $b_{jk} = k/j$ . The detailed calculation of exponents in the continuous approximation of the birth–death process and the Galton–Watson branching process is provided in the following section.

**Moments of Population Density in the Continuous Approximation of the Birth–Death Process and the Galton–Watson Branching Process.** The forward Kolmogorov equation for the continuous approximation of the birth–death process reads (8–10)

$$\frac{\partial p(x,t)}{\partial t} = -\alpha \frac{\partial [xp(x,t)]}{\partial x} + \frac{\beta}{2} \frac{\partial^2 [xp(x,t)]}{\partial x^2}, \quad [\text{S18}]$$

where  $p(x,t)$  is the probability density function for the population density  $x$  at time  $t$  (here,  $x \in \mathbb{R}$  is the population density and should not be confused with the frequency of multiplicative factors used in the rest of the paper). Eq. S18 is the continuous approximation of a birth–death process with birth rate  $\lambda$  and death rate  $\mu$  such that  $\alpha = \lambda - \mu$  and  $\beta = \lambda + \mu$ . Eq. S18 also arises as the continuous approximation of the Galton–Watson branching process for large populations (8–10). The solution of Eq. S18 with initial condition  $x(0) = x_0$  is known (7) and is equal to

$$p(x,t) = \frac{2\alpha}{\beta(e^{at} - 1)} \sqrt{\frac{x_0 e^{at}}{x}} e^{-2\alpha(x_0 e^{at} + x)/\beta(e^{at} - 1)} I_1 \left[ \frac{4\alpha(x_0 x e^{at})^{1/2}}{\beta(e^{at} - 1)} \right], \quad [\text{S19}]$$

where  $I_1$  is the modified Bessel function of the first kind. Differentiation with respect to  $\gamma$  of the identity  $\int_0^\infty dx I_1(x) e^{-\gamma x^2} = e^{1/(4\gamma)} - 1$  gives the equation

$$C \int_0^\infty dx x^k x^{-(1/2)} I_1(x^{1/2} A) e^{Bx} = 2CA^{-(2k+1)} \left( -\frac{d}{d\gamma} \right)^k \Big|_{\gamma = -(B/A^2)} (e^{1/(4\gamma)} - 1), \quad [\text{S20}]$$

which allows calculating the moments of Eq. S19 with  $A = 4\alpha(x_0 e^{at})^{1/2}/\beta(e^{at} - 1)$ ,  $B = 2\alpha/\beta(e^{at} - 1)$ , and  $C = (2\alpha(x_0 e^{at})^{1/2}/\beta(e^{at} - 1)) \exp[-(2\alpha x_0 e^{at}/\beta(e^{at} - 1))]$ . For an expanding population,  $\alpha > 0$ ; thus asymptotically for large  $t$ ,

$$\begin{aligned} A &\propto e^{-(at/2)}, \\ B &\propto e^{-at}, \\ C &\propto e^{-(at/2)}. \end{aligned} \quad [\text{S21}]$$

Therefore,  $\gamma = -(B/A^2)$  tends to a constant and one has

$$\langle x^k \rangle \propto CA^{-2k+1} \propto (e^{at})^k, \quad [\text{S22}]$$

which implies that, asymptotically, the generalized TL holds with exponent  $b_{jk} = k/j$ .

- den Hollander F (2008) *Large Deviations*, Fields Institute Monographs (Am Math Soc, Providence, RI).
- Dembo A, Zeitouni O (2009) *Large Deviations Techniques and Applications*, Stochastic Modelling and Applied Probability (Springer, Berlin Heidelberg).
- Touchette H (2009) The large deviation approach to statistical mechanics. *Phys Rep* 478(1-3):1–69.
- Cohen JE (2014) Stochastic population dynamics in a Markovian environment implies Taylor’s power law of fluctuation scaling. *Theor Popul Biol* 93:30–37.
- Cohen JE, Xu M, Schuster WSF (2013) Stochastic multiplicative population growth predicts and interprets Taylor’s power law of fluctuation scaling. *Proc Biol Sci* 280(1757):20122955.
- Cohen JE (2013) Taylor’s power law of fluctuation scaling and the growth-rate theorem. *Theor Popul Biol* 88:94–100.
- Bailey N (1964) *The Elements of Stochastic Processes* (Wiley, New York).
- Feller W (1957) *An Introduction to Probability Theory and its Applications* (Wiley, New York), Vol 1.
- Harris T (1963) *The Theory of Branching Processes* (Springer, Berlin).
- Rubin K, Pruessner G, Pavliotis G (2014) Mapping multiplicative to additive noise. *J Phys A: Math Theor* 47(19):195001.









**Table S1. Sample exponents for the interspecific generalized TL on carabid beetles abundances in woodland sites, data from ref. 1**

$j, k$	$k/j$	1961		1962		1963		1964		1965		1966	
		$b_{jk} \pm SE$	$R^2$	$b_{jk} \pm SE$	$R^2$	$b_{jk} \pm SE$	$R^2$	$b_{jk} \pm SE$	$R^2$	$b_{jk} \pm SE$	$R^2$	$b_{jk} \pm SE$	$R^2$
1, 2	2	2.03 ± 0.09	0.988	2.07 ± 0.04	0.995	2.00 ± 0.07	0.988	1.96 ± 0.09	0.977	2.01 ± 0.07	0.989	1.97 ± 0.06	0.995
1, 3	3	3.04 ± 0.18	0.976	3.13 ± 0.09	0.991	3.00 ± 0.15	0.977	2.94 ± 0.20	0.957	3.00 ± 0.16	0.976	2.90 ± 0.12	0.989
1, 4	4	4.03 ± 0.28	0.968	4.20 ± 0.14	0.988	4.01 ± 0.23	0.971	3.92 ± 0.29	0.947	4.00 ± 0.24	0.967	3.83 ± 0.18	0.985
No. points		9		13		11		12		11		9	

The column  $k/j$  gives the asymptotic model prediction for the exponent  $b_{jk}$ . The estimates  $b_{jk}$  (mean ± SE) are the least-squares slopes of  $\log(N^k)$  vs.  $\log(N)$ .  $R^2$  is the squared correlation coefficient. Nonlinearity was checked with least-squares quadratic regression on log-log coordinates. The coefficient of the second power term did not differ significantly from 0 in any of the regressions; hence, the null hypothesis of linearity was not rejected.

1. den Boer P (1977) *Dispersal Power and Survival*, Miscellaneous Papers 14 (Landbouwhogeschool Wageningen, Wageningen, The Netherlands).

**Table S2. Sample exponents for the interspecific generalized TL on carabid beetles abundances in heath sites, data from ref. 1**

$j, k$	$k/j$	1963		1964		1965		1966	
		$b_{jk} \pm SE$	$R^2$	$b_{jk} \pm SE$	$R^2$	$b_{jk} \pm SE$	$R^2$	$b_{jk} \pm SE$	$R^2$
1, 2	2	1.99 ± 0.05	0.993	2.02 ± 0.04	0.995	1.98 ± 0.08	0.982	2.02 ± 0.06	0.986
1, 3	3	2.98 ± 0.09	0.987	3.03 ± 0.08	0.990	2.97 ± 0.17	0.965	3.05 ± 0.13	0.974
1, 4	4	3.83 ± 0.18	0.985	3.96 ± 0.14	0.983	4.04 ± 0.12	0.987	3.98 ± 0.26	0.956
No. points		16		16		13		17	

Organized the same as Table S1.

1. den Boer P (1977) *Dispersal Power and Survival*, Miscellaneous Papers 14 (Landbouwhogeschool Wageningen, Wageningen, The Netherlands).

**Table S3. Tests of whether multiplicative growth factors of carabid beetle abundances have the same means and variances over sites**

Test	% of $P < 0.05$	No. species
Complete block $F$	4.3	23
Friedman rank	4.2	24
Kruskal-Wallis	0	24
$K$ sample $T$	0	23
Bartlett	29.6	27
Brown-Forsythe	3.7	27
Conover	7.1	28
Levene	25.9	27

Shown is the percentage of  $P$  values smaller than 0.05 across all species, for several statistical tests. The percentage refers to the number of species used in the test, reported in the third column.

**Table S4. Tests of whether multiplicative growth factors of carabid beetle abundances have the same means and variances over years**

Test	% of $P < 0.05$	No. species
Complete block $F$	14.3	14
Friedman rank	20.0	15
Kruskal–Wallis	53.3	15
$K$ Sample $T$	35.7	14
Bartlett	48.1	27
Brown–Forsythe	7.4	27
Conover	14.3	28
Levene	51.9	27

Shown is the percentage of  $P$  values smaller than 0.05 across all species for several statistical tests. The percentage refers to the number of species used in the test, reported in the third column.