# 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 $X^i(t)=L_t^i(r)$; 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 ($i=1,\ldots,R$). We now define $x^+=\max\{X^1(t),\ldots,X^R(t)\}$ and observe that

$$\frac{1}{t}\log\mathbb{P}\left[X^1(t)>x\right]\le\frac{1}{t}\log\mathbb{P}[x^+>x]\le\frac{1}{t}\log\mathbb{P}\left[X^1(t)>x\right]+\frac{1}{t}\log(R). \quad \textbf{[S1]}$$

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 ($\lim_{t\to\infty}$) in Eq. **S1** and knowing that $L^t(r)$ satisfies a LDP, one has

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

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

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

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

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

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

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

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

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

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 $\chi=\{r_1,\ldots,r_w\}$, where all $r_i$ are strictly positive and at least two $r_i$ are different. We label the state space $\chi=\{1\leftrightarrow r_1,\ldots,w\leftrightarrow r_w\}$. Let the transition matrix $\Pi$ be twofold irreducible (i.e., $\Pi$ irreducible and $\Pi\,\Pi^\top$ irreducible, where $\Pi^\top$ 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)

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

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 $r_i\ne r_j$ for some $i\ne j$ 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 $\mathrm{Var}[N(t)]$ 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 $L_t^i=(L_t^i(r_1),\ldots,L_t^i(r_w))$ and the superscript $i$ indicates the $i$th independent realization of the chain ($i=1,\ldots,R$). We now define $y^+=\max\{Y^1(t),\ldots,Y^R(t)\}$ and observe that, for every $\epsilon>0$,

$$\mathbb{P}(y^+>\epsilon)\le R\,\mathbb{P}\big(Y^1(t)>\epsilon\big). \qquad \textbf{[S8]}$$

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

$$\lim_{t\to\infty}\mathbb{P}(y^+>\epsilon)=0. \qquad \textbf{[S9]}$$

In this context, we can approximate the sample exponent with

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

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

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

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

$$\lim_{t\to\infty}\mathbb{P}(|b_R(t)-2|>\delta)=0. \qquad \textbf{[S12]}$$

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., $\lim_{t\to\infty}(1/t)\log\mathbb{E}[N(t)]=0$. 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, $\lim_{t\to\infty}t^{-1}\log\mathbb{E}[N(t)]=\sup_{x\in[0,1]}[G(x)-I_\Pi(x)]=0$ for $\lambda=\lambda_c$ such that

$$\log\left\{\frac{1}{2}\left[(1-\lambda_c)(r+s)+\sqrt{4(2\lambda_c-1)rs+(\lambda_c-1)^2(r+s)^2}\right]\right\}=0, \qquad \textbf{[S13]}$$

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 $0<r\le 1/s$ and $r\ge 1$ if $s>1$ and for $1\le r\le 1/s$ 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 $\lim_{t\to\infty}t^{-1}\log\mathbb{E}[N(t)^j]=\sup_{x\in[0,1]}[jG(x)-I_\Pi(x)]=0$.

**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. 2*B*), 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. 2*B* exemplifies the above discussion: Initially, the sample TL exponent is different from the theoretical asymptotic prediction for the population TL exponent (Fig. 2*B*, dashed upper horizontal line). For $3\le t\le 12$, the asymptotic prediction for the population TL exponent (Fig. 2*B*, 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 $\lim_{t\to\infty}(1/t)\log\mathbb{E}[N(t)^k]=\sup_{x\in[0,1]}[kG(x)-q(x)x+\xi(q(x))]$. 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 $\lim_{t\to\infty}(1/t)\log\mathbb{E}[N(t)^k]=k\log s+\xi(k\log(r/s))=\log[s^k\zeta(\Pi_{k\log(r/s)})]=\log\zeta(\Pi\,\mathrm{diag}(r,s)^k)$, 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 $\Pi^\top$ that corresponds to the row-stochastic matrix $\Pi$; because $\zeta(\mathrm{diag}(r,s)^k\Pi^\top)=\zeta(\Pi\,\mathrm{diag}(r,s)^k)$, 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)

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

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 $v(0,t) = 0$, $(\partial v/\partial\theta)(\theta,t) = (\lambda-\mu)e^{(\lambda-\mu)t} (e^h / (-e^h\lambda + \mu)^2) \propto e^{(\lambda-\mu)t}$, 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

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

where the lowercase-$o$ notation indicates that the remaining terms are negligible in the limit $t \to \infty$. 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,

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

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

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 $t \to \infty$. From Eq. **S15** it follows that

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

The asymptotic behavior of exponents $\lim_{t\to\infty}(1/t)\log\langle N^k\rangle = k(\lambda-\mu)$ 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)

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

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

$$p(x,t) = \frac{2\alpha}{\beta(e^{\alpha t}-1)}\sqrt{\frac{x_0e^{\alpha t}}{x}}\,e^{-2\alpha(x_0e^{\alpha t}+x)/\beta(e^{\alpha t}-1)}I_1\left[\frac{4\alpha(x_0xe^{\alpha t})^{1/2}}{\beta(e^{\alpha t}-1)}\right], \quad \textbf{[S19]}$$

where $I_1$ is the modified Bessel function of the first kind. Differentiation with respect to $\gamma$ of the identity $\int_0^\infty dxI_1(x)e^{-\gamma x^2} = e^{1/(4\gamma)}-1$ gives the equation

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

which allows calculating the moments of Eq. **S19** with $A = 4\alpha(x_0e^{\alpha t})^{1/2}/\beta(e^{\alpha t}-1)$, $B=2\alpha/\beta(e^{\alpha t}-1)$, and $C=(2\alpha(x_0e^{\alpha t})^{1/2}/\beta(e^{\alpha t}-1))\exp[-(2\alpha x_0e^{\alpha t}/\beta(e^{\alpha t}-1))]$. For an expanding population, $\alpha>0$; thus asymptotically for large $t$,

$$A\propto e^{-(\alpha t/2)},\quad B\propto e^{-\alpha t},\quad C\propto e^{-(\alpha t/2)}. \quad \textbf{[S21]}$$

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

$$\left\langle x^k\right\rangle \propto CA^{-2k+1}\propto\left(e^{\alpha t}\right)^k, \quad \textbf{[S22]}$$

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


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

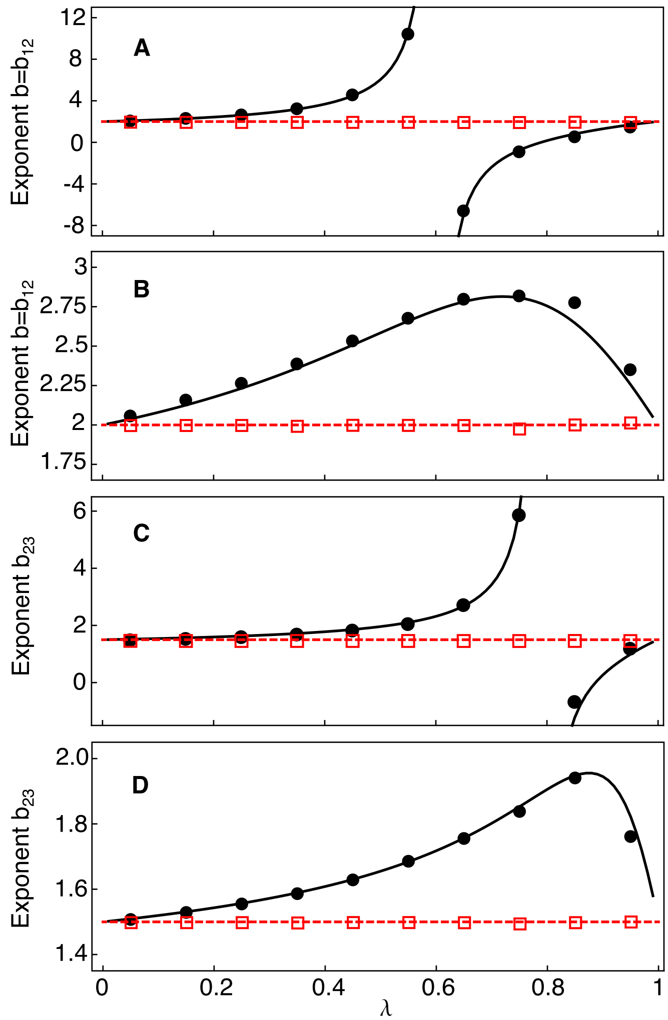


**Fig. S1.** TL exponent $b = b_{12}$ and generalized exponent $b_{23}$ for different values of the transition probability $\lambda$ (*A* and *B* are as in Fig. 1). The sample exponents computed in simulations of a two-state multiplicative process with symmetric transition matrix in the two regimes $1 \ll t \ll \log R$ (black solid circles, $R = 10^6$ 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. **S6**) and sample (red dashed line, $b = b_{12} = 2$ and $b_{23} = 3/2$) exponents. In the simulations, the sample exponent $b = b_{12}$ was computed by least-squares fitting of $\log \mathrm{Var}[N(t)]$ as a function of $\log \mathbb{E}[N(t)]$ for the last 6 (black circles) and 200 (red squares) time steps. The sample exponent $b_{23}$ was computed by least-squares fitting of $\log \mathbb{E}[N(t)^3]$ as a function of $\log \mathbb{E}[N(t)^2]$ in the same fashion. In *A*, which has the plotted theoretical result from ref. 1, and *C*, $\chi = \{r, s\} = \{2, 1/4\}$ ($b = b_{12}$ and $b_{23}$ display discontinuities); in *B* and *D*, $\chi = \{r, s\} = \{4, 1/2\}$ (in such a case, $b_{12}$ and $b_{23}$ display no discontinuities).


1. Cohen JE (2014) Taylor's law and abrupt biotic change in a smoothly changing environment. *Theor Ecol* 7(1):77–86.

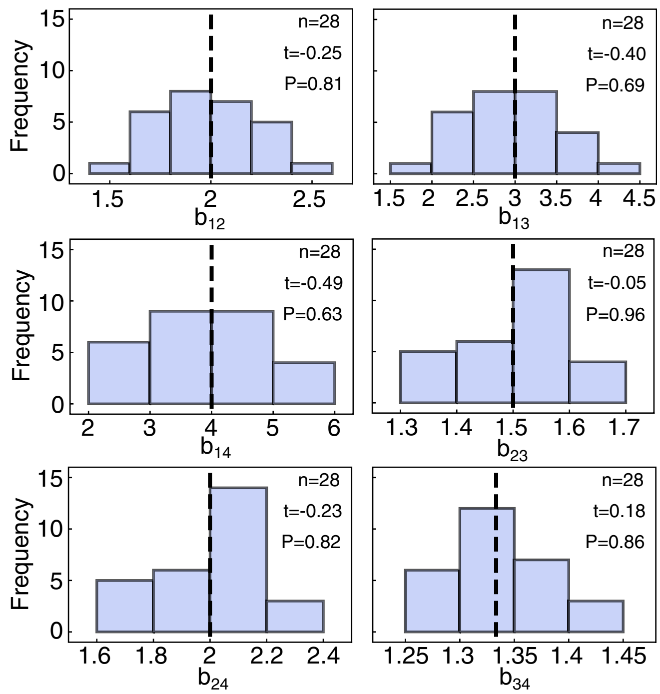


**Fig. S2.** Frequency histogram for the exponent $b_{jk}$ in the intraspecific generalized TL $\langle N^k \rangle = a \langle N^j \rangle^{b_{jk}}$, computed for each species [carabid beetles (1)] across similar sites (woodland or heath). The dashed black line shows the value of the exponent $b_{jk} = k/j$ as the asymptotic model predicted. The binning of data points is determined by using Scott's rule (2). Shown in each panel are the number of observations *n* of $b_{jk}$, the test statistic for the *t* test of the null hypothesis that the sample mean of the values of $b_{jk}$ did not differ significantly from the theoretically predicted mean $k/j$, and the corresponding *P* value.

1. den Boer P (1977) *Dispersal Power and Survival*, Miscellaneous Papers 14 (Landbouwhogeschool Wageningen, Wageningen, The Netherlands).
2. Scott D (1979) On optimal and data-based histograms. *Biometrika* 66(3):605–610.

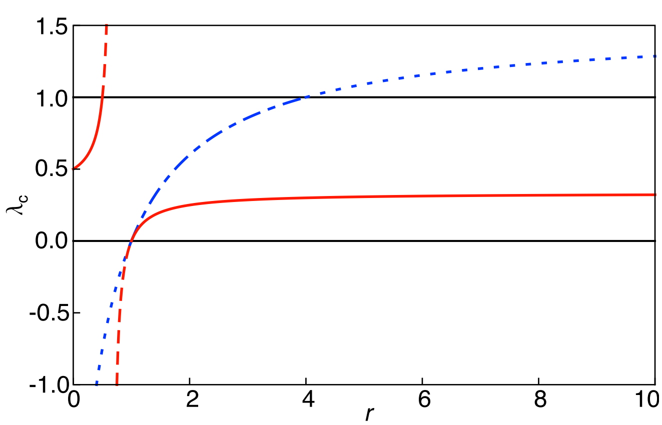


**Fig. S3.** The critical transition probability $\lambda_c$ as a function of *r* (with *s* fixed). Below the black horizontal line at $\lambda_c = 0$ and above the black horizontal line at $\lambda_c = 1$, $\lambda_c$ does not exist. The red (solid for $0 \le \lambda_c \le 1$ and dashed otherwise) and blue (dash-dotted for $0 \le \lambda_c \le 1$ and dotted otherwise) lines $\lambda_c = (1 - r - s + rs)/(-r - s + 2rs)$ were calculated by solving Eq. **S13** with respect to $\lambda_c$ with, respectively, $s = 2$ and $s = 1/4$. For any given *s*, $\lambda_c = 0$ for $r = 1$ and $\lambda_c = 1$ for $r = 1/s$.

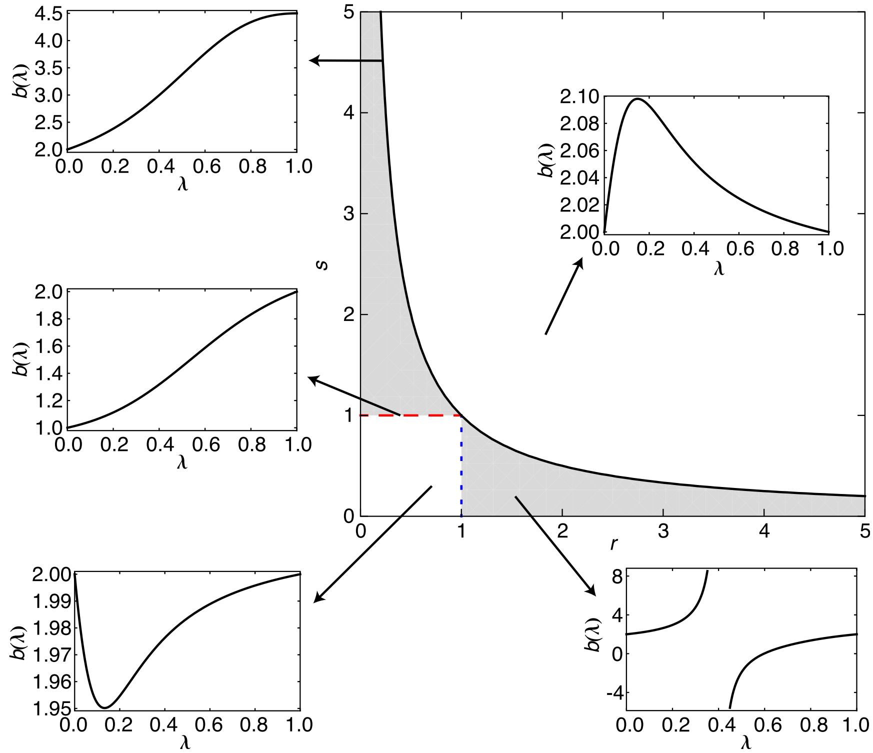


**Fig. S4.** Existence of a critical transition probability $\lambda_c$. Small panels show the population exponent $b(\lambda)$ (Eq. **9**) for various choices of the multiplicative factors in different regions of the plane $(r, s)$ (large panel). Only in the interior of the gray region of the plane $(r, s)$, $\lambda_c$ exists. The solid black line represents the curve $rs = 1$.

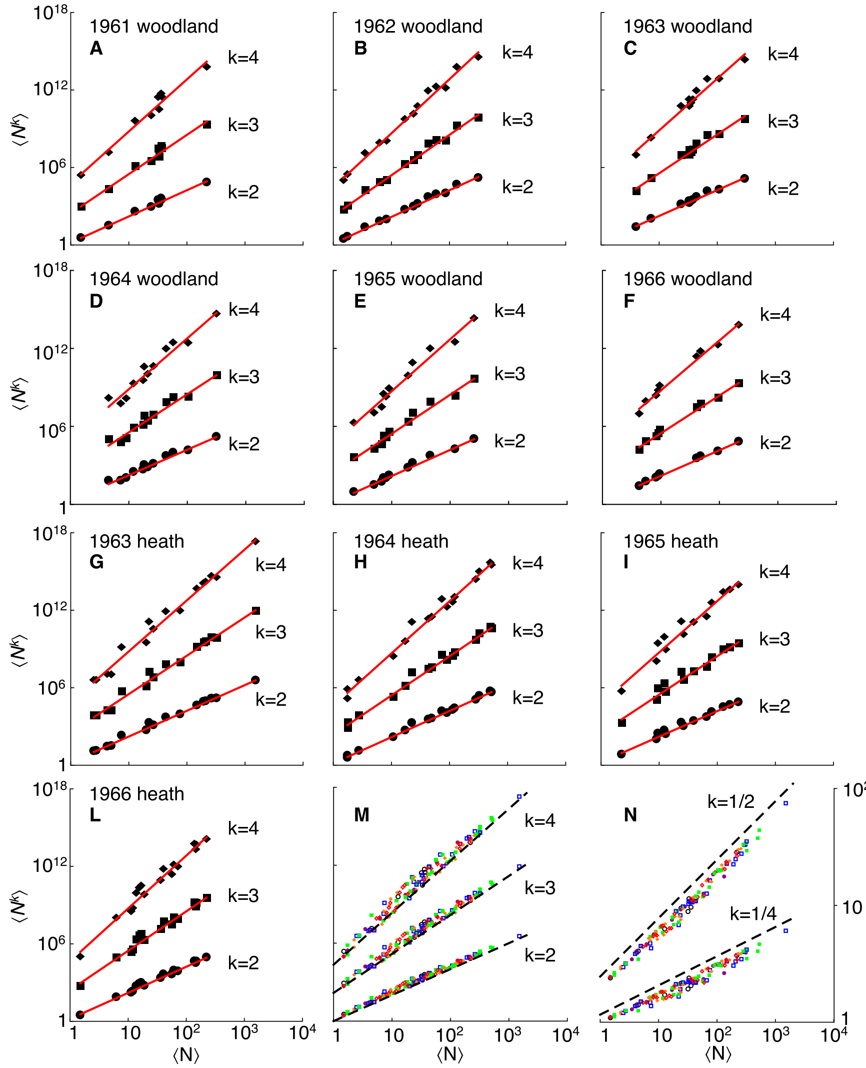


**Fig. S5.** Generalized TL for interspecific patterns of abundance of carabid beetles, data from ref. 1. *(A–L)* Double logarithmic plots of $\langle N^k \rangle$ vs. $\langle N \rangle$ for all species in separate years and site type (black symbols). The red lines show the least-squares regressions of $\log\langle N^k \rangle$ vs. $\log\langle N \rangle$ (Tables S1 and S2). Offsets are introduced in the data and in the linear regressions to aid visual inspection. *(M* and *N)* Double logarithmic plot of $\langle N^k \rangle$ vs. $\langle N \rangle$ for all species, years, and site type, with integer (*M*) and noninteger (*N*) *k*. Each data point refers to sample moments computed for 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). The color and symbol code does not distinguish site type. Dashed black lines of slope $b_{1k} = k/1 = k$ (asymptotic model prediction for the sample exponent) and arbitrary intercept are shown in each plot. Offsets are introduced in the data to aid visual inspection.

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

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

| | | 1961 | | 1962 | | 1963 | | 1964 | | 1965 | | 1966 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $j, k$ | $k/j$ | $b_{jk} \pm$ 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 $\log\langle N^k \rangle$ vs. $\log\langle N \rangle$. $R^2$ 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**

| | | 1963 | | 1964 | | 1965 | | 1966 | |
|---|---|---|---|---|---|---|---|---|---|
| $j, k$ | $k/j$ | $b_{jk} \pm$ 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.