

Greedy algorithms and Zipf laws

José Moran¹ and Jean-Philippe Bouchaud²

¹*Centre d'Analyse et de Mathématiques Sociales,
EHESS, 54 Boulevard Raspail, 75006 Paris**

²*Capital Fund Management, 23 Rue de l'Université, 75007 Paris*

(Dated: March 1, 2018)

Abstract

We consider a simple model of firm/city/etc. growth based on a multi-item criterion: whenever entity B fares better than entity A on a subset of M items out of K , the agent originally in A moves to B. We solve the model analytically in the cases $K = 1$ and $K \rightarrow \infty$. The resulting stationary distribution of sizes is generically a Zipf-law provided $M > K/2$. When $M \leq K/2$, no selection occurs and the size distribution remains thin-tailed. In the special case $M = K$, one needs to regularise the problem by introducing a small “default” probability ϕ . We find that the stationary distribution has a power-law tail that becomes a Zipf-law when $\phi \rightarrow 0$. The approach to the stationary state can also be characterized, with strong similarities with a simple “aging” model considered by Barrat & Mézard.

INTRODUCTION

Among all power-law distributions, the Zipf law appears to play a special role. Mathematically, the Zipf law corresponds to a marginally (logarithmically) divergent first moment – distributions decaying faster have a finite mean, while distributions decaying slower have an infinite mean. Empirically, this law appears to hold in a variety of situations, from the frequency of words to the node degree of the Internet [1, 2], the size of companies [3] or the population of cities [4, 5]; for reviews see [5, 6]. Correspondingly, there has been a flurry of possible explanation for the ubiquitous Zipf law. One standard explanation for the Zipf law is based on stochastic multiplicative growth, which under relatively mild assumptions (weak redistribution between entities and/or total mass conservation) can be shown to converge towards a Zipf distribution [4, 5, 7]. More recently, Gualdi & Mandel [8] and Axtell [9] have proposed a different mechanism in the context of company sizes, based on the idea that larger firms are more “attractive” than smaller firms, and individual agents tend to switch from small firms to large firms.

Inspired by these ideas, we propose here a schematic model where agents randomly select over time better and better items, possibly using multiple criteria. For definiteness, we will think about firms, but other interpretations are possible. We postulate that each firm i is characterized by a set of K attributes, with scores $\mathbf{x}^i \in [0, 1]^K$ that all agents agree on. For instance, x_1^i is the score given to the wages paid by firm i , x_2^i the score given to its location, x_3^i the score given to its work environment, etc. Firm j is deemed “better” than firm i when the number of attributes for which j has a higher score than i exceeds M (with $1 \leq M \leq K$). In the single criterion case $K = 1$, we find that the stationary distribution of firm sizes is a Zipf distribution when the reset rate tends to zero. Other power-laws are possible, depending on the specific search mechanism. In the multi-criterion case, we show that the asymptotic distribution is a truncated Zipf law as soon as $M > K/2$ (up to logarithmic corrections when $K > 1$), with a diverging cutoff in the limit $K \rightarrow \infty$, $M/K > 1/2$. For $M \leq K/2$, there is no “condensation” and hence no power-law tails, as expected.

THE SINGLE CRITERION MODEL

Greedy algorithm

We first consider the case $K = 1$. Each firm $i = 1, \dots, N$ is characterized by a single score $x^i \in [0, 1]$. Firms are populated with agents; the fraction of agents belonging to firm i at time t is denoted $n_i(t)$. At each time infinitesimal step dt , an agent is chosen at random with probability Γdt ; s/he picks at random a firm among the ones that are “better” than the firm s/he presently works for, and moves to the newly chosen firm. We consider from the outset the limit $N \rightarrow \infty$ and assume without loss of generality that the x^i are distributed uniformly in $[0, 1]$. (One can always transform any chosen distribution of x into the uniform one by a change of variable that preserves the ordering of the scores.) We also choose to label firms by their score x^i rather than by i , and describe the system at time t by the function $n(x, t)$ such that the fraction of agents belonging to firms with score between x and $x + dx$ is $n(x, t)dx$. The evolution of $n(x, t)$ when agents use the greedy algorithm above is, for $x < 1$:

$$\frac{\partial n(x, t)}{\partial t} = -\Gamma n(x, t) + \Gamma \int_0^x dy \frac{n(y, t)}{1 - y}, \quad (1)$$

where the second term in the RHS corresponds to agents in firms with score $y < x$ choosing their next firm in the interval $[y, 1]$ with uniform probability. As written, Eq. (1) obviously leads to a stationary state where all agents condense in the firm with the highest score (i.e. $x = 1$ for $N \rightarrow \infty$.) Some regularization is needed to make the problem non-trivial and interesting. One realistic assumption is that each firm has a probability φ per unit time to go under, in which case all its employees find new jobs with uniform probability in the remaining firms. The defaulted firm is replaced by a new one with an x chosen uniformly in $[0, 1]$. This leads to the following modified Master equation

$$\frac{\partial n(x, t)}{\partial t} = -\Gamma n(x, t) + \Gamma \int_0^x dy \frac{n(y, t)}{1 - y} - \varphi n(x, t) + \varphi. \quad (2)$$

Another, nearly equivalent interpretation, is to assume that new agents are injected in a growing economy, with a constant rate φ and a uniform “deposition rate” among all firms.

We first look for a stationary solution to Eq. (2). Setting the LHS to zero and changing variable to $u = 1 - x$, the stationary solution $n_{\text{st}}(u)$ must obey the following differential

equation:

$$-\Gamma n'_{\text{st}}(u) - \Gamma \frac{n_{\text{st}}(u)}{u} - \varphi n'_{\text{st}}(u) = 0. \quad (3)$$

Introducing $\phi := \varphi/\Gamma$, the normalized solution to the above equation reads:

$$n_{\text{st}}(u) = \frac{\phi}{1+\phi} \frac{1}{u^{\frac{1}{1+\phi}}}. \quad (4)$$

This equation gives us the equilibrium fraction of agents in firms of quality $x = 1 - u$.

We are now in position to derive the stationary distribution of the size s of firms. Since the distribution of u 's is uniform, one has

$$P(s) = \int_0^1 du \delta\left(s - \frac{\phi}{1+\phi} \frac{1}{u^{\frac{1}{1+\phi}}}\right) = \phi \left(\frac{\phi}{1+\phi}\right)^{1+\phi} s^{-2-\phi}, \quad (s > \phi/(1+\phi)) \quad (5)$$

which tends to a Zipf law when $\phi \rightarrow 0$. For companies, Axtell's results translate to $\phi \approx 0.06$. If one takes $\varphi = 0.07/\text{year}$ as in [10], this suggests a reasonable search frequency of $\Gamma \approx 1/\text{year}$, although empirical data suggests a median job tenure of about 4 years in the United States [11].

Note that our model is a continuous formulation of the SSR process studied by Corominas-Murtra et al. in [12]. It is also closely related to the well known Simon model for growing networks [13], where each new site links proportionally to the size of pre-existing clusters with probability $1/(1+\phi)$ and creates a new singleton cluster with probability $\phi/(1+\phi)$.

One can in fact obtain the full time dependent solution of Eq. (2) as follows: we first write $n(u, t) = n_{\text{st}}(u) + m(u, t)e^{-\varphi t}$ with $\int du m(u, t) = 0$. Substituting in Eq. (2) one finds that $m(u, t)$ obeys exactly the un-regularized equation (1). The physical interpretation of this equation is very simple: it describes the multiplicative process where the random variable $u(t + dt)$ is equal to u_t with probability $1 - \Gamma dt$, and to $a_t u_t$ with probability Γdt , where a_t is a random variable with uniform distribution in $[0, 1]$. Hence, the quantity $\ell := -\log(u)$ is an additive random walk, for which the probability distribution evolves according to a standard Fokker-Planck equation:

$$\frac{\partial P(\ell, t)}{\partial t} = -\Gamma \frac{\partial P(\ell, t)}{\partial \ell} + \frac{\Gamma}{2} \frac{\partial^2 P(\ell, t)}{\partial \ell^2} \quad (6)$$

Hence:

$$m(u, t) = \frac{1}{u} \int d\xi \frac{M_0(\xi)}{\sqrt{2\pi\Gamma t}} \exp\left(-\frac{(\log u + \xi + \Gamma t)^2}{2\Gamma t}\right), \quad (7)$$

where $M_0(\xi)$ is the initial condition, such that $um(u, 0) = M_0(-\log u)$.

Greedy but Myopic

In the model above, we assume that each agent preselects “better” firms and makes a move with probability 1 if chosen. Assume now that if chosen our agent picks a firm at random among all possible firms, and decides to move only if the chosen firm is better than the current one. An agent in a firm of quality $x = 1 - u$ will thus only move with probability u . The corresponding Master equation now reads:

$$\frac{\partial n(x, t)}{\partial t} = -\Gamma n(x, t)(1 - x) + \Gamma \int_0^x dy n(y, t) - \varphi n(x, t) + \varphi. \quad (8)$$

Its stationary solution now reads:

$$n_{\text{st}}(u) = \frac{\phi(\phi + 1)}{(u + \phi)^2}, \quad (9)$$

and the stationary distribution of firm sizes is:

$$P(s) = \int_0^1 du \delta\left(s - \frac{\phi(\phi + 1)}{(u + \phi)^2}\right) = \frac{\sqrt{\phi(1 + \phi)}}{2} s^{-3/2} \quad (10)$$

where the last equality holds in the range $s \in [\phi/(1 + \phi), 1 + 1/\phi]$, outside which $P(s)$ is zero. Hence we find that in this model, the tail of the distribution is *fatter* than that of the Zipf law. This (perhaps counter-intuitive) result is due to the fact that the time spent by each agent in large firms is longer in the present setting, due to the myopic search algorithm that becomes extremely inefficient when $u \rightarrow 0$. [14]

The full dynamics of the model can also be solved. Using Laplace transforms and the method of characteristics, the deviation from equilibrium is given by:

$$m(u, t)e^{-\varphi t} = e^{-(u+\phi)\Gamma t} m(u, 0) + \Gamma t e^{-(u+\phi)\Gamma t} \int_u^1 dv m(v, 0), \quad (11)$$

indicating that the system converges to the stationary state $n_{\text{st}}(u)$ after a time $\sim (u\Gamma)^{-1}$ for $u \gg \phi^{-1}$ and after a time $\sim \varphi^{-1}$ for small u 's. Hence, ϕ^{-1} acts as both a cut-off limiting the values attainable by s , and as the relaxation time-scale of the system.

When $\varphi = 0$, Eq. (8) is in fact equivalent to a model studied by Barrat and Mézard in [15], where a particle moves in a landscape of randomly distributed energy levels. At zero temperature, the energy can only go down and the Master equation of the Barrat and Mézard model reads

$$\partial_t p(E, t) = -p(E, t) \int_{-\infty}^E dE' \mathcal{P}(E') + \int_E^{\infty} dE' \mathcal{P}(E') p(E', t) \quad (12)$$

where $p(E, t)$ is the probability to find the system at energy E at time t and $\mathcal{P}(E)$ is the density of energy states. Eq. (12) is equivalent to our model through the mapping $u = \int_{-\infty}^E dE' \mathcal{P}(E')$. In terms of u , the exact dynamical solution is given by:

$$n(u, t) = e^{-\Gamma u t} n(u, 0) + \Gamma t e^{-u \Gamma t} \int_u^1 du n(u, 0). \quad (13)$$

This solution becomes universal at large times in the regime where $u = z/\Gamma t$, z finite. One finds:

$$n(u, t) = \Gamma t \mathcal{F}(u \Gamma t); \quad \mathcal{F}(z) = e^{-z}. \quad (14)$$

We also compute the distribution $p_t(\tau)$ of trapping times $\tau := \frac{1}{\Gamma u}$ at time t , given by:

$$p_t(\tau) = \frac{t}{\tau^2} e^{-\frac{t}{\tau}} \quad (15)$$

in full agreement with the findings of [15] in the appropriate limit $t \rightarrow \infty$, $\tau \rightarrow \infty$ with $t/\tau = O(1)$. In our context, the trapping time represents the time spent by an agent in a firm before moving to another firm.

Interpolating between models

The two models above can be cast in a single framework, where agents spend a u -dependent time $\tau(u)$ in their firm before looking for a better firm. The general Master equation then reads

$$\frac{\partial n(u, t)}{\partial t} = -\frac{n(u, t)}{\tau(u)} + \int_u^1 dv \frac{n(v, t)}{v \tau(v)} - \varphi n(u, t) + \varphi. \quad (16)$$

The case $\tau(u) = \Gamma^{-1}$ corresponds to the Greedy model introduced previously, while the Greedy but myopic model is when $\tau(u) = (\Gamma u)^{-1}$. One can interpolate between the two models by choosing $\tau(u) = \Gamma^{-1} u^{-\beta}$. The stationary solution is then found to be, for $\beta \neq 0$,

$$n_{\text{st}}(u) = \frac{Z_\beta(\phi)}{(u^\beta + \phi)^{1+\frac{1}{\beta}}}, \quad (17)$$

where Z_β is a normalisation, and the solution given in Eq. (4) for $\beta = 0$. The corresponding tail of firm size distribution behaves as $P(s) \sim s^{-\frac{\beta+2}{\beta+1}}$, recovering the Zipf law for $\beta = 0$ and the $s^{-3/2}$ behaviour for $\beta = 1$. Note that the above result continues to make sense as long as $\beta > -1$.

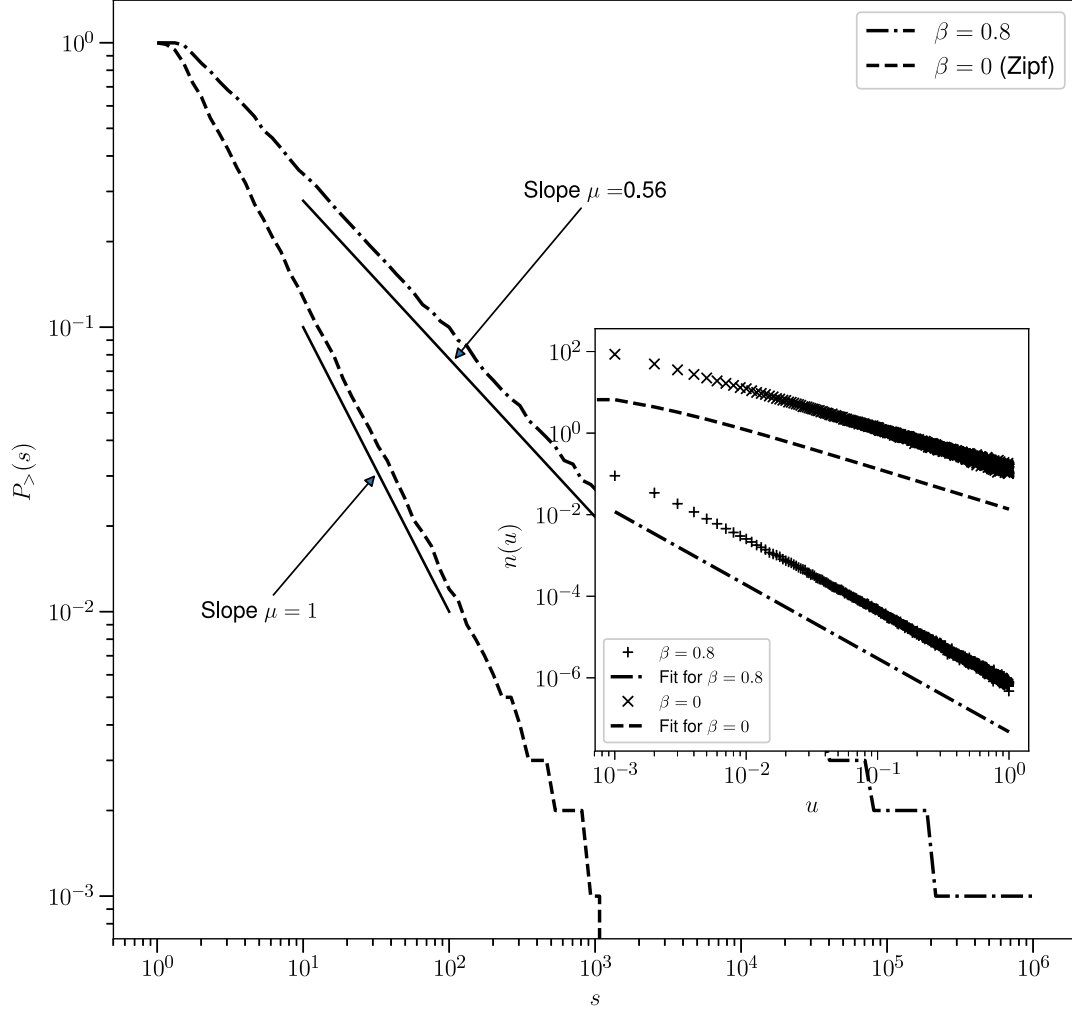


FIG. 1. Log-log plots of $P_>(s) := \int_s^\infty ds' P(s') \sim s^{-\mu}$, with $\mu := \frac{1}{\beta+1}$ and $n_{\text{st}}(u)$ as given by numerical simulations along with their corresponding fits to the analytical results.

THE MULTI-CRITERION CASE

Let us formally define the model by specifying the rules according to which an agent in a firm characterized by a vector of scores $\vec{x} = (x_1, \dots, x_K)$ can move into a new firm with scores $\vec{y} = (y_1, \dots, y_K)$. Setting again $u_i = 1 - x_i$ and $v_i = 1 - y_i$, the transition is allowed

when there exists a subset $(i_1, \dots, i_M) \in \{1, \dots, K\}$ such that:

$$\begin{aligned} v_{i_1} &< u_{i_1} \\ &\vdots \\ v_{i_M} &< u_{i_M}. \end{aligned} \tag{18}$$

If this is the case, the corresponding transition matrix $W_{\vec{u} \rightarrow \vec{v}}$ is set to unity, otherwise $W_{\vec{u} \rightarrow \vec{v}} = 0$.

Defining $n(\vec{u}, t)$ as the density of agents in firms of score \vec{u} at time t , the general Master equation we will consider is

$$\frac{\partial n(\vec{u}, t)}{\partial t} = -\Gamma n(\vec{u}, t) \omega_{K,M}^\beta(\vec{u}) + \Gamma \int_{W_{\vec{v} \rightarrow \vec{u}} \neq 0} d\vec{v} n(\vec{v}, t) \omega_{K,M}^{\beta-1}(\vec{u}) - \varphi n(\vec{u}, t) + \varphi, \tag{19}$$

where

$$\omega_{K,M}(\vec{u}) = \int_{[0,1]^K} d\vec{v} W_{\vec{u} \rightarrow \vec{v}} \tag{20}$$

is the total volume accessible for \vec{u} . The case $\beta = 0$ corresponds to the “greedy” algorithm of the previous section, while $\beta = 1$ corresponds to its “myopic” counterpart.

In the stationary state, one can map models with different values of β in the limit $\varphi \rightarrow 0$ by writing:

$$n_{\text{st}, \beta \neq 0}(\vec{u}) \propto \frac{n_{\text{st}, \beta=0}(\vec{u})}{\omega_{K,M}(\vec{u})^\beta}. \tag{21}$$

Grand Slam

Let us first consider the case $M = K$, i.e. when all scores must improve for the agent to change firm. In this case, the available volume $\omega_{K,K}(\vec{u})$ is easy to compute and reads:

$$\omega_{K,K}(\vec{u}) = \prod_{i=1}^K u_i. \tag{22}$$

The stationary distribution then takes a factorized form, which for $\beta = 0$ reads:

$$n_{\text{st}}(\vec{u}) = \prod_{i=1}^K \frac{\phi}{1 + \phi} \frac{1}{u_i^{\frac{1}{1+\phi}}}. \tag{23}$$

This result is expected, because each of the K scores, in this setting, follows an independent one-dimensional greedy process.

The stationary firm size distribution is given, for $\beta = 0$, by

$$P(s) = \int_{[0,1]^K} d\vec{u} \delta \left(s - \prod_{i=1}^K \frac{\phi}{1+\phi} \frac{1}{u_i^{\frac{1}{1+\phi}}} \right). \quad (24)$$

This can be computed by first introducing the following function:

$$F_K(t) = \int_{[0,1]^K} d\vec{u} \delta \left(t - \prod_{i=1}^K u_i \right). \quad (25)$$

Integrating over – say – u_K , one easily derives the following recursion relation:

$$F_K(t) = \int_t^1 ds \frac{F_{K-1}(s)}{s}, \quad (26)$$

which solves as

$$F_K(t) = \frac{(\log t^{-1})^{K-1}}{(K-1)!}. \quad (27)$$

Injecting this expression into Eq. (24) we thus find:

$$P(s) = \int dt F_K(t) \delta \left(s - A_K t^{-\frac{1}{1+\phi}} \right), \quad (28)$$

with $A_K = \phi^K / (1 + \phi)^K$. The final result is a Zipf law with a K dependent logarithmic correction:

$$P(s \rightarrow \infty) \propto \left(\log \frac{s}{A_K} \right)^{K-1} s^{-2-\phi}. \quad (29)$$

Marginal distributions

In the following, we consider $M < K$. As we will see, no regularization is needed in this case since agents can always “escape” from good firms. Hence, we set $\varphi = 0$ henceforth. Taking the distribution of the new scores \vec{v} conditioned on the previous scores \vec{u} , and integrating over (v_1, \dots, v_{K-1}) allows us to compute the marginal distribution of v_K , $P(v_K)$, given \vec{u}

$$\begin{aligned} P(v_K | \vec{u}) &= \frac{\theta(u_K - v_K) \omega_{K-1, M-1}(\vec{u}_{\bar{K}}) + \theta(v_K - u_K) \omega_{K-1, M}(\vec{u}_{\bar{K}})}{u_K \omega_{K-1, M-1}(\vec{u}_{\bar{K}}) + (1 - u_K) \omega_{K-1, M}(\vec{u}_{\bar{K}})} \\ &= \frac{\theta(u_K - v_K) + \nu(\vec{u}_{\bar{K}}) \theta(v_K - u_K)}{u_K + \nu(\vec{u}_{\bar{K}})(1 - u_K)}, \end{aligned} \quad (30)$$

where $\theta(x) = 1$ for $x > 0$ and zero otherwise, $\vec{u}_{\bar{K}} = (u_1, \dots, u_{K-1})$ and

$$\nu(\vec{u}_{\bar{K}}) := \frac{\omega_{K-1, M}}{\omega_{K-1, M-1}} \quad (31)$$

where $\nu = \frac{\omega_{K-1,M}(\vec{u}_{\bar{K}})}{\omega_{K-1,M-1}(\vec{u}_{\bar{K}})}$.

Assuming that ν converges to a constant value ν^* (which is in fact only true in the limit $K \rightarrow \infty$, see below), the stationary marginal distribution of scores $P_{\text{st}}(v)$ must verify

$$P_{\text{st}}(v) = \int_v^1 du \frac{P_{\text{st}}(u)}{u + \nu^*(1-u)} + \nu^* \int_0^v du \frac{P_{\text{st}}(u)}{u + \nu^*(1-u)} \quad (32)$$

which has a (normalized) solution:

$$P_{\text{st}}(u) = \frac{\nu^* - 1}{\log(\nu^*)} \frac{1}{u + (1 - \nu^*)u} \quad (33)$$

We now show that in the large K limit, ν^* can be self-consistently computed and only depends on the ratio $\alpha = \frac{M}{K}$.

The large K limit

We first note that the volume $\omega_{K,M}(\vec{u})$ can be represented exactly using binary spin variables $\sigma_k = \pm 1$, as

$$\begin{aligned} \omega_{K,M}(\vec{u}) &= \int_{[0,1]^K} d\vec{v} W_{\vec{u} \rightarrow \vec{v}} \\ &= \sum_{\{\sigma_k = \pm 1\}} \prod_{k=1}^K u_k^{\frac{1+\sigma_k}{2}} (1-u_k)^{\frac{1-\sigma_k}{2}} \theta \left(\sum_{k=1}^K \sigma_k - (2M - K) \right). \end{aligned} \quad (34)$$

Using the Fourier representation of the Dirac delta function, one finds:

$$\omega_{K,M}(\vec{u}) \propto \int_{2\alpha-1}^1 d\mu \int d\lambda e^{-i\mu\lambda} \sum_{\{\sigma_k = \pm 1\}} \exp \left(\sum_{k=1}^K (h(u_k) + i\lambda) \sigma_k + \frac{\log(u_k(1-u_k))}{2} \right) \quad (35)$$

where $h(u) = \frac{1}{2} \log \left(\frac{u}{1-u} \right)$.

Now we are left with computing the partition function of a system of Ising spins with local fields $h(u_k) + i\lambda$. Summing over all configurations $\{\sigma_k = \pm 1\}$ we find that each coordinate contributes to the partition function with a term $(\cosh(h(u_k) + i\lambda))$. After a few algebraic manipulations we obtain

$$\omega_{K,M}(\vec{u}) \propto \int_{2\alpha-1}^1 d\mu \int d\lambda \exp \left(\sum_{k=1}^K (\log(u_k + e^{-2i\lambda}(1-u_k)) - i\lambda(\mu - 1)) \right) \quad (36)$$

In the large K limit two simplifications occur. First, one can replace the sum $\sum_{k=1}^K f(u_k)$ with the integral $K \int_0^1 du P(u) f(u)$ when f is an arbitrary regular function. Second, one can

calculate the integral over λ using a saddle-point method. Hence:

$$\omega_{K,M}(\vec{u}) \propto \int_{2\alpha-1}^1 d\mu \exp K \left(\int_0^1 du P(u) (\log(u + e^{-2\lambda^*(\mu)}(1-u)) - \lambda^*(\mu)(\mu-1)) \right) \quad (37)$$

where $\lambda^*(\mu)$ satisfies the saddle-point equation:

$$\mu = \int_0^1 du P(u) \frac{\tanh(\lambda^*) + 2u - 1}{1 + \tanh(\lambda^*)(2u - 1)}. \quad (38)$$

Assuming that $P(u)$ has reached its stationary limit $P_{\text{st}}(u)$ given by Eq. (33), we therefore find:

$$\mu = \begin{cases} \frac{2(\nu^*-1)\varepsilon \log(\varepsilon) - (\varepsilon-1) \log(\nu^*)(\nu^*+\varepsilon)}{(\varepsilon-1) \log(\nu^*)(\varepsilon-\nu^*)} & \text{if } \varepsilon \neq \nu^* \\ \frac{2}{\log(\nu^*)} - \frac{\nu^*+1}{\nu^*-1} & \text{if } \varepsilon = \nu^* \end{cases} \quad (39)$$

with $\varepsilon := e^{-2\lambda^*}$. This last equation must be solved for λ^* for a given pair of values μ, ν^* . The remaining integral over μ is again estimated using a second saddle-point, leading to:

$$\begin{aligned} \mu^*(\alpha, \nu^*) &= \operatorname{argmax}_{\mu \in [2\alpha-1, 1]} \left(\int_0^1 du P_{\text{st}}(u) (\log(u + e^{-2\lambda^*(\mu)}(1-u)) - \lambda^*(\mu)(\mu-1)) \right) \\ &:= \operatorname{argmax}_{\mu \in [2\alpha-1, 1]} g(\mu, \nu^*) \end{aligned} \quad (40)$$

and we obtain our final result for $\omega_{K,M}(\vec{u})$:

$$\omega_{K,M}(\vec{u}) \propto \exp \left(K \int_0^1 du P_{\text{st}}(u) (\log(u + e^{-2\lambda^*(\mu^*)}(1-u)) - \lambda^*(\mu^*)(\mu^*-1)) \right) \quad (41)$$

One can check that for all $\nu \in [0, 1]$, the function $g(\mu)$ is decreasing in the interval $[0, 1]$, and has a maximum for $\mu = 0$ with $g(0) = 0$ as depicted in Fig. 2 for $\nu = \varepsilon$ or $\nu = 1$. Assuming that the saddle point is $\mu^* = 2\alpha - 1$ when $\alpha > \frac{1}{2}$ and $\mu^* = 0$ otherwise, one can compute the relevant volume ratio for large K and for $\alpha > \frac{1}{2}$ using Eq. (41):

$$\frac{\omega_{K-1,M}}{\omega_{K-1,M}} = \frac{\omega_{K-1,\alpha(1+\frac{1}{K})}}{\omega_{K-1,\alpha+\frac{\alpha-1}{K}}} \approx e^{-2\lambda^*(\mu^*)} := \varepsilon^*. \quad (42)$$

validating thus our assumption for the saddle-point. Setting $\nu^* = \varepsilon^*$ in Eq. (39), we finally obtain an implicit expression for ν^* :

$$\begin{aligned} 1 - 2\alpha &= \frac{\nu^* + 1}{\nu^* - 1} - \frac{2}{\log(\nu^*)} \text{ when } \alpha > \frac{1}{2} \\ \nu^* &= 1 \text{ when } \alpha \leq \frac{1}{2}. \end{aligned} \quad (43)$$

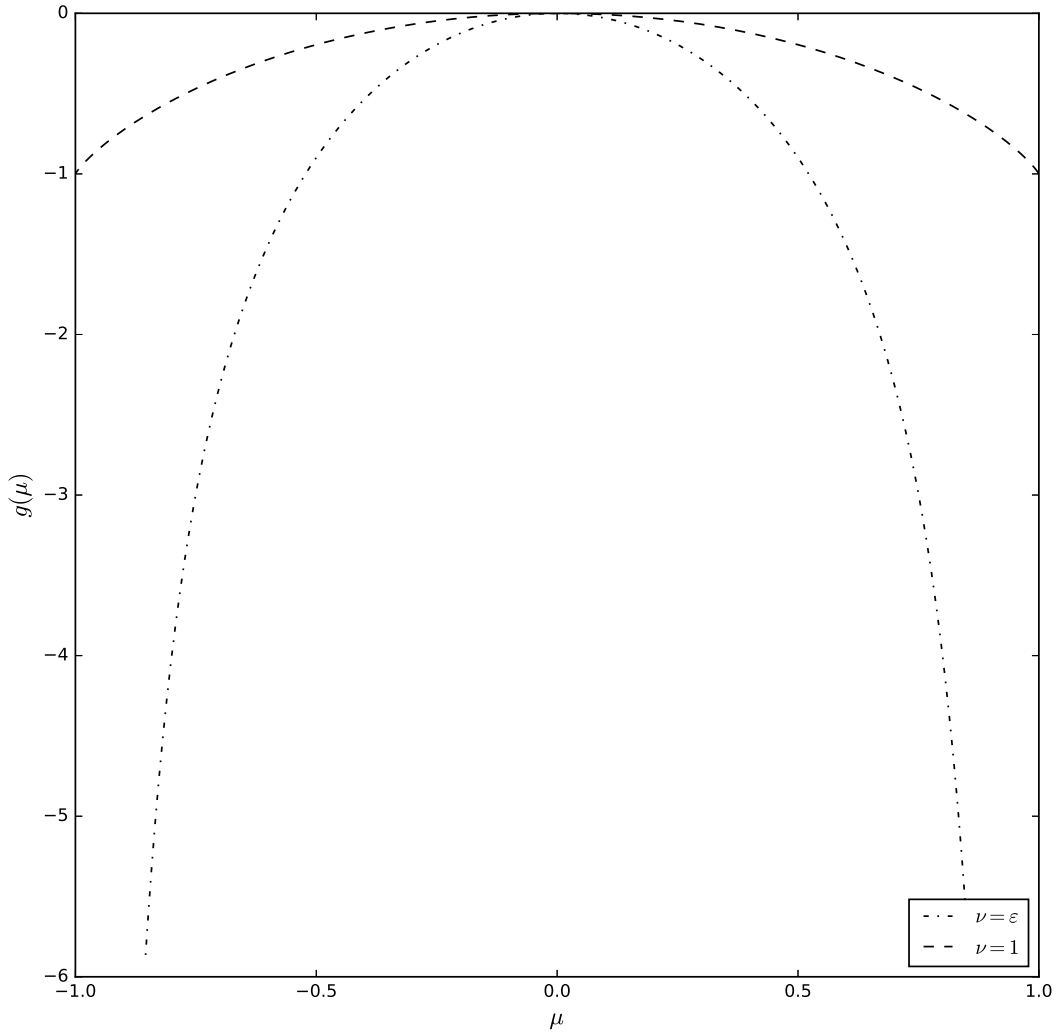


FIG. 2. Plot of $g(\mu)$ as computed for $\nu = 1$ and $\nu = \varepsilon$. Note that the maximum is reached in both cases for $g(0) = 0$, allowing one to establish the result in Eq. (43).

We compare our analytical results with numerical simulations in Figs. 3 and 4. The shape of $P_{\text{st}}(u)$ describes very well the numerical results even for small values of K . The value of ν^* , however, is only slowly reached when $K \rightarrow \infty$.

When $\alpha \rightarrow 1$ (i.e. when $M \rightarrow K$), the solution for ν^* is exponentially small:

$$\nu^* \approx e^{-\frac{1}{1-\alpha}}, \quad (44)$$

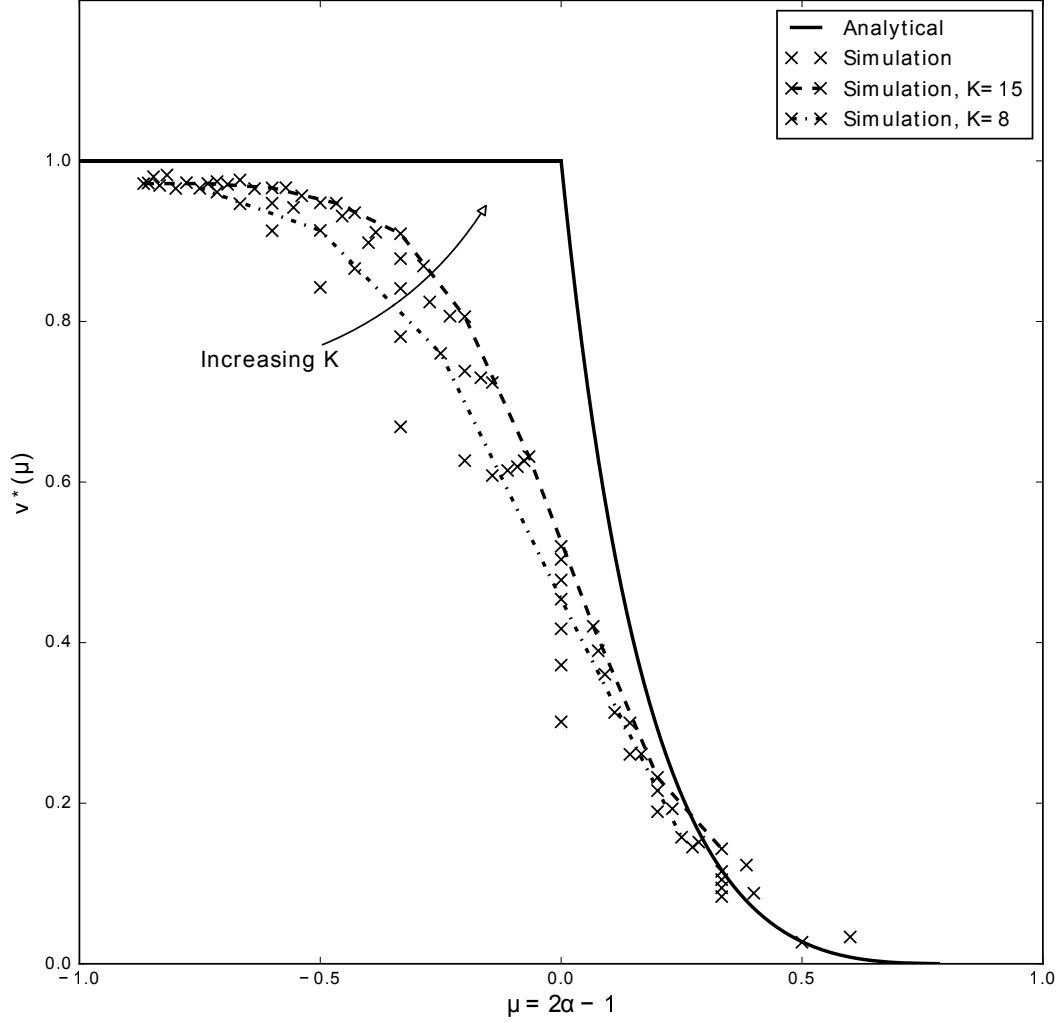


FIG. 3. Plot of ν^* computed using the empirical transition matrix from simulations with $K = 2 \dots 15$ vs. the analytical value given in Eq. (43). Note that the empirical values get closer to the analytical curve as K gets larger, as shown by the dashed lines over the values for $K = 15$ and $K = 8$.

whereas when $\alpha \rightarrow \frac{1}{2}$ from above, $\nu^* \approx 1 - 6(2\alpha - 1)$ and tends to 1. When $\alpha < \frac{1}{2}$, $\nu^* = 1$ and $P_{\text{st}}(u) = 1$. Hence, in the case where $M < K/2$, the distribution of scores is uniform in the large time limit: the selection process is too weak to induce any differentiation between firms.

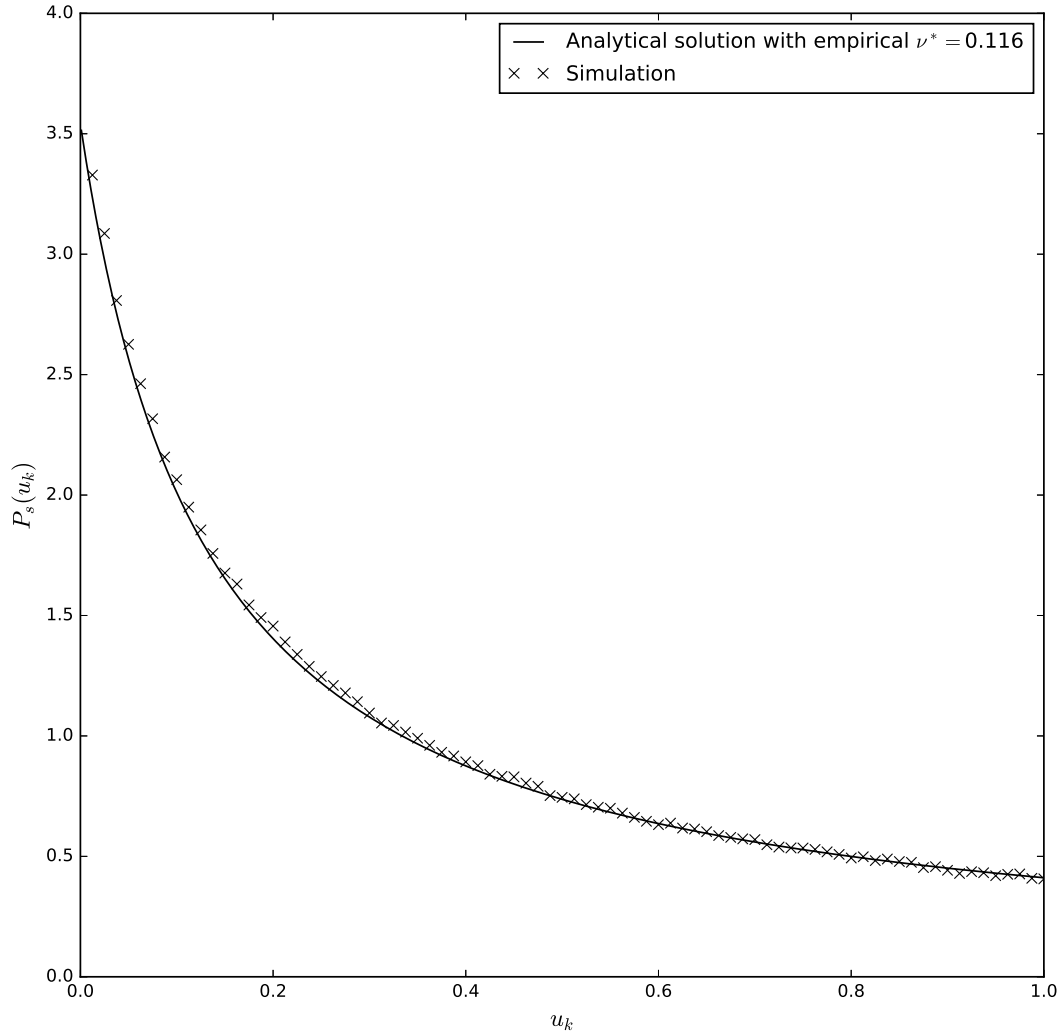


FIG. 4. Plot of empirical and analytical $P_{\text{st}}(u)$ for $M = 8$ and $K = 12$. The value of ν^* was again determined using the empirical transition matrix and plugged into Eq. (33).

Since in the large K limit the dynamics of scores decouple, the full stationary distribution $n_{\text{st}}(\vec{u})$ is simply given by

$$n_{\text{st}}(\vec{u}) \propto \prod_{i=1}^K \frac{1}{(u_i + (1 - u_i)\nu^*)^{1+\beta}} \quad (45)$$

This is the central result of this section. Notice that the re-injection mechanism that we introduced in the mono-criterion case is not needed here since $\nu^* > 0$ whenever $\alpha < 1$.

When $\alpha \rightarrow 1$, $\nu^* \rightarrow 0$ and the stationary distribution is extremely close to the result for $M = K$, see Eq. (23) for $\phi = 0$.

Now the distribution of firm sizes can be computed as

$$\begin{aligned} P(s) &= \int_{[\nu;1]^K} d\vec{v} \, \delta \left(s - A_K \prod_{i=1}^K v_i \right) \\ &= \int_{[\nu^K;1]} dt \, \delta(n - A_K t^{-1}) F_K(t, \nu) \end{aligned} \quad (46)$$

with $v_i = u_i + (1 - u_i)\nu^*$, $A_K = \left(\frac{\nu^* - 1}{\log(\nu^*)} \right)^K$. We have introduced the function $F_K(t, \nu) = \int_{[\nu;1]^K} d\vec{u} \, \delta(t - \prod u_i)$ that satisfies:

$$\begin{cases} F_1(t, \nu) = \theta(t - \nu) \\ F_K(t, \nu) = \int_t^{\min(\frac{t}{\nu}, 1)} \frac{ds}{s} F_{K-1}(s, \nu), \end{cases} \quad (47)$$

which can be computed explicitly for small values of K : [16]

$$\begin{aligned} F_2(t, \nu) &= \theta(t - \nu) \log(t^{-1}) + \theta(\nu - t) \theta(t - \nu^2) \log\left(\frac{t}{\nu^2}\right) \\ 2F_3(t, \nu) &= \theta(t - \nu) \log^2(t^{-1}) + \theta(\nu - t) \theta(t - \nu^2) \left(\log\left(\frac{\nu}{t}\right)^2 + \log\left(\frac{t}{\nu^2}\right)^2 \right) \\ &\quad + \theta(\nu^2 - t) \theta(t - \nu^3) \log^2\left(\frac{t}{\nu^3}\right). \end{aligned} \quad (48)$$

One can convince one self that this structure generalises for larger K s, with contributions $\frac{\log^{K-1}(\frac{\nu^j}{t})}{(K-1)!}$ and $\frac{\log^{K-1}(\frac{t}{\nu^{j+1}})}{(K-1)!}$ when $t \in [\nu^{j+1}; \nu^j]$, $j = 1 \dots K-1$. In the end intervals one finds $F_K(t, \nu) = \frac{(\log^{K-1}(t^{-1}))}{(K-1)!}$ and $F_K(t) = \frac{\log^{K-1}(\frac{t}{\nu^K})}{(K-1)!}$ for $\nu^K < t < \nu^{K-1}$. Finally, $F_K(t, \nu) = 0$ for $t < \nu^K$. Although quite complicated, one therefore finds that $F_K(t, \nu)$ is a piece-wise polynomial function of $\log t^{-1}$ in the whole interval. Hence, the distribution $P(s)$, given by:

$$P(s) \propto \begin{cases} s^{-2} F_K\left(\frac{A_K}{s}, \nu\right) & \text{for } s < s^* = \frac{A_K}{\nu^{*K}} \\ 0 & \text{for } s > s^* = \frac{A_K}{\nu^{*K}}, \end{cases} \quad (49)$$

is a Zipf-law modulated by powers of logarithms in a broad interval that extends up to $s^* \underset{\nu \rightarrow 0}{\simeq} \left(\nu \log\left(\frac{1}{\nu}\right) \right)^{-K}$. The important point here is that as soon as $\nu^* < 1$, the upper cut-off s^* becomes very large when K is large, all the more so when α is close to unity. An example of a plot of $P(s)$ is given in Fig. 5.

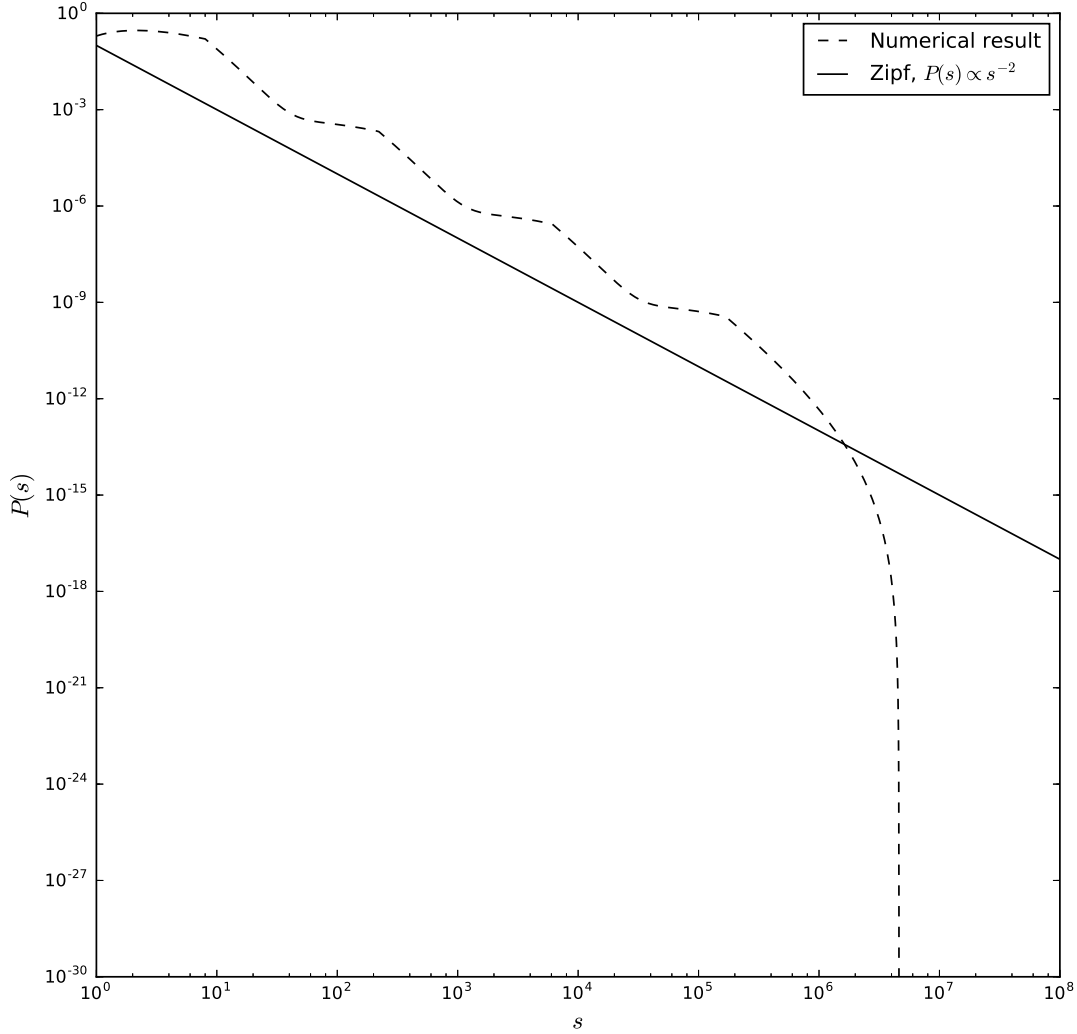


FIG. 5. Log-log plot of $P(s)$ for $K = 5$ and $M = 4$, compared to the Zipf law s^{-2} . The figure was plotted using the numerical value of $\nu^* \simeq 0.036$. Note the modulation introduced by the logarithmic corrections encoded in the function $F_K(u, \nu)$.

CONCLUSION

We have introduced a simple model of firm/city/etc. growth based on the idea that agents choose to switch from entity A to entity B depending on a multi-item criterion: whenever entity B fares better than entity A on a subset of M items out of K , the agent originally

in A moves to B. We have solved the model analytically in the cases $K = 1$ and $K \rightarrow \infty$. The resulting stationary distribution of sizes is generically a Zipf-law provided $M > K/2$. When $M \leq K/2$, no selection occurs and the size distribution remains thin-tailed. In the special case $M = K$, some regularisation process must be introduced to prevent the whole system from condensating into the “best” entity. Introducing a small probability ϕ that each entity defaults and redistributes all its agents in surviving entities, one finds that the stationary distribution has a power-law tail that becomes a Zipf-law when $\phi \rightarrow 0$. The approach to the stationary state has also been characterized, with strong similarities with a simple “aging” model considered by Barrat & Mézard [15]. Although our model for $K = 1$ looks superficially similar to those considered by Gualdi & Mandel [8] and Axtell [9], we have not been able to elicit a precise mapping between these models, and more generally with the slew of other mechanisms that lead to Zipf-laws.

Acknowledgements: We benefited from very fruitful discussions with M. Cordi, S. Gualdi, M. Tarzia and F. Zamponi, who participated actively to the first stages of this project.

* jose.moran@polytechnique.org

- [1] L. A. Adamic and B. Huberman, *Glottometrics* **3** (2001).
- [2] B. A. Huberman and L. A. Adamic, *Condensed Matter* (1999), 10.2139/ssrn.166108.
- [3] R. Axtell, *Science* **293**, 1818 (2001), <http://science.sciencemag.org/content/293/5536/1818.full.pdf>.
- [4] X. Gabaix, *The Quarterly Journal of Economics* **114**, 739 (1999).
- [5] A. I. Saichev, Y. Malevergne, and D. Sornette, *Theory of Zipf’s law and beyond*, Vol. 632 (Springer Science & Business Media, 2009).
- [6] X. Gabaix, *Annu. Rev. Econ.* **1**, 255 (2009).
- [7] J.-P. Bouchaud and M. Mézard, *Physica A: Statistical Mechanics and its Applications* **282**, 536 (2000).
- [8] S. Gualdi and A. Mandel, *Journal of Economic Dynamics and Control* **73**, 61 (2016).
- [9] R. Axtell, Manuscript, George Mason University (2013).
- [10] M. I. G. Daep, M. J. Hamilton, G. B. West, and L. M. A. Bettencourt, *Journal of The Royal Society Interface* **12** (2015), 10.1098/rsif.2015.0120, <http://rsif.royalsocietypublishing.org/content/12/106/20150120.full.pdf>.
- [11] U. S. D. o. L. Bureau of Labor Statistics, “Employee tenure in 2016,” (2016), consulted Feb.

2018.

- [12] B. Corominas-Murtra, R. Hanel, and S. Thurner, Nature Scientific Reports **7** (2017), 10.1038/s41598-017-098
- [13] H. A. Simon, Biometrika **42**, 425 (1955).
- [14] The reader may realize that the difference between the two models is precisely related to the difference between naive Monte-Carlo and the so-called Bortz-Kalos-Lebowitz algorithm.
- [15] A. Barrat and M. Mézard, Journal de Physique I **5**, 941 (1995).
- [16] Expressed in terms of $\ell_i = \log(u_i + (1 - u_i)\nu^*)$, the problem can be mapped onto a sum of independent random variables, uniformly distributed in $[0, 1]$. The corresponding distribution for $K > 1$ is called the Irvin-Hall distribution.