Continuum rich-get-richer processes: Mean field analysis with an application to firm size

David Rushing Dewhurst,[1,2,*] Christopher M. Danforth,[2,†] and Peter Sheridan Dodds[2,‡]

[1]*The MITRE Corporation, McLean, Virginia 22102, USA*

[2]*Department of Mathematics & Statistics, Vermont Complex Systems Center, Computational Story Laboratory,*
& the Vermont Advanced Computing Core, The University of Vermont, Burlington, Vermont 05401, USA

Classical rich-get-richer models have found much success in being able to broadly reproduce the statistics and dynamics of diverse real complex systems. These rich-get-richer models are based on classical urn models and unfold step by step in discrete time. Here, we consider a natural variation acting on a temporal continuum in the form of a partial differential equation (PDE). We first show that the continuum version of Simon's canonical preferential attachment model exhibits an identical size distribution. In relaxing Simon's assumption of a linear growth mechanism, we consider the case of an arbitrary growth kernel and find the general solution to the resultant PDE. We then extend the PDE to multiple spatial dimensions, again determining the general solution. We then relax the zero-diffusion assumption and find an envelope of solutions to the general model in the presence of small fluctuations. Finally, we apply the model to size and wealth distributions of firms. We obtain power-law scaling for both to be concordant with simulations as well as observational data, providing a parsimonious theoretical explanation for these phenomena.

I. INTRODUCTION

In 1955, Simon described a general version of a rich-get-richer process that generates power-law size distributions $P(s) \sim s^{-\gamma}$ with a scaling exponent $\gamma > 2$ [1]. Simon's process was adapted by Price to capture the statistics of growing networks, and was later paralleled by the Barabási-Albert model which introduced scale-free networks [2]. Simon's model efficiently captures the statistical properties of a wide variety of real-world phenomena, such as the linking dynamics of the Web [3] and the growth of software distributions [4]. Recently, the present authors and others have shown Simon's model also exhibits a potentially pronounced first-mover advantage and that this feature may be consistent with the growth of real systems [5].

In Sec. II, we first realize Simon's model in a continuum setting and describe its dynamics for a number of growth kernels. Sections II A and II B describe the continuum version of the model and formulate its analytical solution. In Sec. II C we determine analytically how the size distribution generated by the process is dependent on the growth kernel, and can be proportional not only to any power-law distribution with finite mean ($\gamma > 2$), but also specific instances of the extreme value distribution, while in Sec. II D we analyze the model's behavior when extended to many dimensions. We apply the model to the dynamics of a market economy in Sec. III, showing that the power-law distribution of firms observed empirically and in simulation can be derived from first principles of microeconomic theory with a minimum of assumptions [6].

II. MODEL AND ANALYSIS

We describe Simon's discrete model by means of an economic example. Suppose an individual creates a new firm in some product space with themselves as the sole employee. An individual that enters the product space at time step t must choose between starting a new firm themselves with probability ρ, and choosing to join an existing firm with probability $1 - \rho$ from one of the existing firms, with the likelihood of choosing any particular firm from which to purchase proportional to the number of employees k. We will denote the number of firms of size k at time t by $N_{k,t}$. The general discrete model thus takes the form of the recurrence relation [1]

$$\langle N_{k,t+1} - N_{k,t} \rangle = (1 - \rho)\left(-\frac{k}{t}N_{k,t} + \frac{k-1}{t}N_{k-1,t}\right), \quad (1)$$

where we formalize ρ as an innovation probability. The solution to (1) scales as

$$N_{k,t} \sim t k^{-\gamma}, \quad (2)$$

with $\gamma = 1 + \frac{1}{1-\rho}$. When $\rho \to 0$, the size exponent $\gamma \to 2$, so that the distribution thus obtained borders on infinite mean. Zipf's law for rank-frequency distributions, written $s_r \propto r^{-\alpha}$, is recovered from (2) by setting the Zipf exponent $\alpha = \frac{1}{\gamma-1} = 1 - \rho$ [1]. The corresponding equation for the size of the nth arriving group $S_{n,t}$ is then given by [5]

$$S_{n,t} = \begin{cases} \frac{1}{\Gamma(2-\rho)}\left[\frac{1}{t}\right]^{-(1-\rho)} & \text{if } n = 1, \\ \rho^{1-\rho}\left[\frac{n-1}{t}\right]^{-(1-\rho)} & \text{if } n \geqslant 2. \end{cases} \quad (3)$$

A. From discrete to continuous

Though the discrete model accurately models the size distribution resulting from many real-world processes, it has a number of shortcomings when applied to economic situations.

---
*david.dewhurst@uvm.edu
†chris.danforth@uvm.edu
‡peter.dodds@uvm.edu

Models that exploit mathematical properties of the preferences of a representative agent often perform poorly in the task of explaining economic phenomena such as inequality and skewed wealth distributions [7,8]. Abstraction of these considerations is thus desirable in order to account for idiosyncrasies present at the individual consumer level; the resulting model is a mean field equation and better describes macrobehavior, analogous to the use of deterministic equations of statistical mechanics to describe stochastic interactions among many particles. Moving from discrete to continuous time is sensible as it corresponds better with our notion of reality (people do not make decisions in synchrony at each tick of a universal clock). From a mathematical viewpoint, the resulting equation will be more easily analyzed as a partial differential equation instead of a coupled differential-difference equation. Further, where Simon's model assumes that agents aggregate to firms with growth kernel $r(x) = x$, we drop this assumption and write the growth kernel as some function $r(x)$ to allow for generalization of choice [9]. Finally, we allow the innovation rate ρ to vary in time as $\rho(t)$, which may more realistically capture the process of technological innovation inherent in the present economic system. (We note that Simon considered a time-varying innovation probability in Ref. [1].)

While other authors have considered models that provided important contributions to the understanding of rich-get-richer processes and are *prima facie* similar to ours [9,10], our model differs substantially from those previously created. Previous continuum models have focused solely on networks, while ours is intentionally more general, allowing us to construct parsimonious models of economic phenomena, for example. In addition, these previous models did not treat the most general problem of arbitrary growth kernel $r(x)$ and innovation rate $\rho(t)$ in the manner considered here. In addition, the continuum formulation extended to an arbitrary (finite) number of dimensions is entirely different, as far as we are aware; this formalism can be used in modeling interacting preferential attachment processes in the fields of biology, economics, or sociology.

B. General asymptotic model

Applying the above adjustments to Simon's discrete model, Eq. (1) becomes a boundary value problem for the function determining the intertemporal distribution of firms of size x, written $f(x,t)$,

$$\frac{\partial f}{\partial t} = -\frac{1-\rho(t)}{t}\frac{\partial}{\partial x}[r(x)f], \qquad (4)$$

with the semi-infinite boundary condition $\lim_{x\to\infty} f(x,t) = 0$. (We treat only the case of asymptotic solutions; the question of differing initial distributions of firms is not considered.) We first consider ρ_∞ as a long-run constant innovation rate satisfying $\lim_{t\to\infty}\frac{\rho(t)}{\rho_\infty} = 1$ and solve Eq. (4) by separation of variables. Setting $f(x,t) = X(x)T(t)$, we solve the equations

$$\frac{dT}{dt} \simeq \frac{\lambda}{1-\rho_\infty}\frac{1-\rho(t)}{t}T(t), \qquad (5)$$

$$\frac{dX}{dx} = -\left[\frac{\lambda}{(1-\rho_\infty)r(x)} + \frac{r'(x)}{r(x)}\right]X(x), \qquad (6)$$

where λ is a constant of separation. The general solution to (4) is thus

$$f(x,t) = \frac{c}{r(x)}\exp\left[\frac{\lambda}{1-\rho_\infty}g(x,t)\right], \qquad (7)$$

where $g(x,t) = \int^t \frac{1-\rho(t')}{t'}dt' - \int^x \frac{dx'}{r(x')}$.

Dropping the assumption that $\rho(t) \to \rho_\infty$, we solve (4) in all generality using the method of characteristics. We write the Lagrange-Charpit equations that describe its solution on the characteristic curves as

$$dt = \frac{t\,dx}{[1-\rho(t)]r(x)} = -\frac{t\,df}{[1-\rho(t)]r'(x)f(x,t)}, \qquad (8)$$

and solve the resulting equations

$$\frac{dx}{dt} = \frac{1-\rho(t)}{t}r(x), \qquad (9)$$

$$\frac{df}{dx} = -\frac{r'(x)}{r(x)}f(x,t). \qquad (10)$$

The solution to the first is given implicitly by $\int \frac{dx}{r(x)} + A = \int \frac{1-\rho(t)}{t}dt$, while the solution to the second is $f(x,t) = \frac{B}{r(x)}$. Letting $B = F(A)$ gives the firm density

$$f(x,t) = \frac{1}{r(x)}F\left(\int \frac{1-\rho(t)}{t}dt - \int \frac{dx}{r(x)}\right). \qquad (11)$$

We see that Eq. (7) has the same form as Eq. (11), with $F(\cdot) = \exp(\cdot)$.

C. Asymptotics for example growth kernels

We wish to characterize the long-run behavior of Eq. (7). Recovering the original Simon model is possible by setting $r(x) = x$ and letting the innovation rate remain constant at ρ_∞. Equation (7) then becomes

$$f(x,t) \propto \frac{1}{x}\exp\left[\ln t - \frac{\lambda}{1-\rho_\infty}\int^x \frac{dx'}{x'}\right]$$

$$= tx^{-(1+\frac{\lambda}{1-\rho_\infty})}, \qquad (12)$$

with $\lambda \to 1$ in the long-run time limit. Figure 1 shows solutions of Eq. (4) with the Simon growth kernel $r(x) = x$ and a constant innovation rates $\rho \in \{0.01, 0.1, 0.5\}$. Note that these solutions are pure power laws and thus are linear in log-log space. The exponent $\gamma = 1 + \frac{1}{1-\rho_\infty}$ is the expression found by Simon in Ref. [1]. Any discretization of this continuum process will have a size-rank distribution $S(r) \propto r^{-\alpha}$ with Zipf exponent $\alpha = \frac{1}{\gamma-1} = 1 - \rho_\infty$; the dynamics of the Simon process are thus completely recovered in this case.

For an affine growth kernel $r(x) = a_0 + a_1 x$, the solution is similar,

$$f(x,t) \propto \frac{1}{a_0 + a_1 x}\exp\left[\ln t - \frac{\lambda}{1-\rho_\infty}\int^x \frac{dx'}{a_0 + a_1 x'}\right]$$

$$= t(a_0 + a_1 x)^{-1-\frac{\lambda}{a_1(1-\rho_\infty)}}$$

$$\sim tx^{-(1+\frac{1}{a_1(1-\rho_\infty)})}, \qquad (13)$$

again in the long-run time limit. A general linear growth factor can thus be chosen to result in a power-law distribution with any $\gamma > 1$; as a_1 grows large, $\frac{1}{a_1(1-\rho_\infty)} \to 0$.
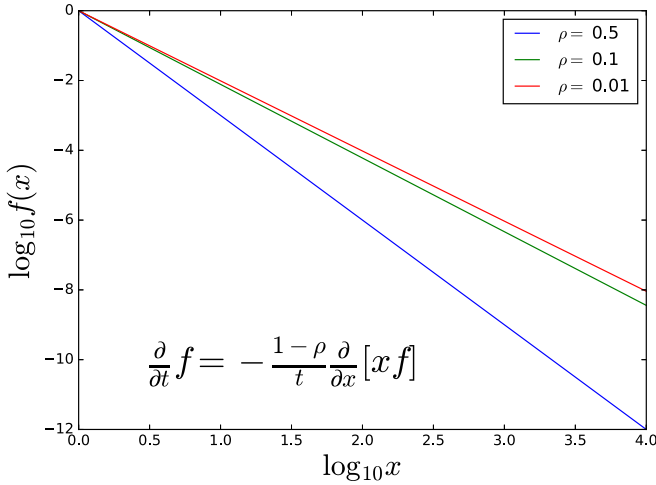
FIG. 1. Solutions to Eq. (4) with growth kernel $r(x) = x$ and innovation rate $\rho(t) = \rho_\infty$.

Considering a monomial power growth factor $r(x) = x^\eta$ ($\eta \neq 1$), we obtain

$$f(x,t) \propto \frac{1}{x^\eta} \exp\left[\ln t - \frac{1}{1-\rho} \int^x \frac{dx'}{x'^\eta}\right]$$
$$= t x^{-\eta} \exp\left[-\frac{x^{-(\eta-1)}}{(1-\rho)[-(\eta-1)]}\right]. \qquad (14)$$

Equation (14) is proportional to a Fréchet distribution for $\eta > 1$ or Weibull distribution for $0 < \eta < 1$. Krapivsky *et al.* found a similar result for n_k, the number of nodes of degree k, in growing random networks [11].

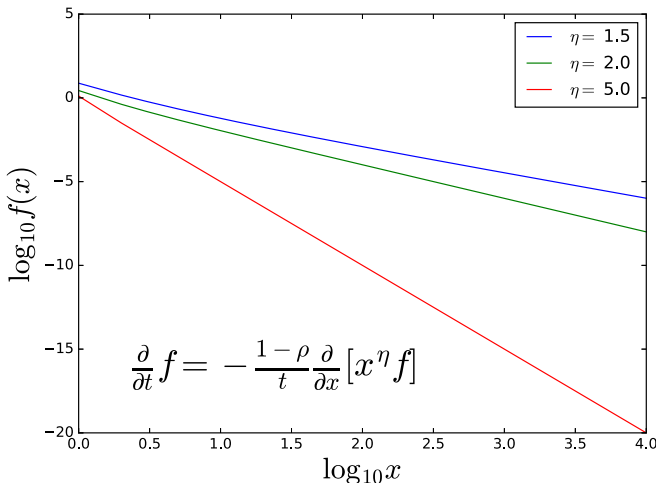Figure 2 shows solutions of Eq. (4) with $r(x) = x^\eta$ for $\eta > 1$.

FIG. 2. Solutions to Eq. (4) with growth kernel $r(x) = x^\eta$ for $\eta > 1$ and innovation rate $\rho(t) = \rho_\infty$.

D. Preferential attachment in many dimensions

We now extend the model, Eq. (4), to multiple dimensions. The general model reads

$$\frac{\partial f}{\partial t} = -\frac{1-\rho(t)}{t} \mathcal{L}_{\text{div}}(r(x_1,\ldots,x_N), f)$$
$$+ \mathcal{L}_{\text{fluc}}(D(x_1,\ldots,x_N)f), \qquad (15)$$

where $r(x_1,\ldots,x_N)$ is a general growth kernel, $\mathcal{L}_{\text{div}}$ is a linear differential operator containing a divergence term, the generator of the rich-get-richer process, and $\mathcal{L}_{\text{fluc}}$ is an operator containing information on the random fluctuations of f. We describe several models and solve one of them analytically.

The most straightforward generalization of Eq. (4) holds constant the assumption of zero fluctuations and has a generator given by the divergence of the vector field $\mathbf{E} = \sum_{i=1}^N [r_k(x_k)f]\mathbf{e}_k$, with $\mathbf{e}_k$ the standard orthonormal basis for the particular vector space under study. In the case of Cartesian coordinates, the generator becomes

$$\mathcal{L}_{\text{div}} = \sum_{i=1}^N \frac{\partial}{\partial x_k}[r_k(x_k)f].$$

This case has an analytical solution given below in Sec. II D 1 that directly parallels the results given in Sec. II B. When $r_k = r_k(x_1,\ldots,x_N)$, the resulting equation is not separable and resists analytical solution.

A reasonable generalization relaxes the zero-diffusion assumption to study the process under the influence of small perturbations. Assuming small random fluctuations gives a time-dependent Fokker-Planck equation for f which we now describe. Letting $\mathbf{D} = \sum_{i,j} D_{ij}(x_1,\ldots,x_N)$ be the covariance matrix, we have (in Cartesian coordinates)

$$\frac{\partial f}{\partial t} = -\frac{1-\rho(t)}{t}\nabla \cdot \mathbf{E} + H[\mathbf{D}f], \qquad (16)$$

where $H = \sum_{i,j} \frac{\partial^2}{\partial x_j \partial x_i}$ is the diffusion operator. This equation is also, in general, not possible to solve analytically. However, an envelope of solutions is available in any small time interval, as shown in Sec. II D 2.

1. Separable growth kernels r_k

Maintaining the assumption of a steady-state constant innovation rate ρ_∞, the equation governing the distribution of firms as a function of time and N spatial variables $x_1,\ldots,x_N$ is

$$\frac{\partial f}{\partial t} = -\frac{1-\rho(t)}{t}\sum_{k=1}^N \frac{\partial}{\partial x_k}(r_k f), \qquad (17)$$

where $r_k = r_k(x_k)$. This equation is again separable with the solution given by

$$f(x_1,\ldots,x_N,t) = T(t)\prod_{k=1}^N X_k(x_k).$$

Substituting the above into Eq. (17) gives

$$\frac{\partial\left[T(t)\prod_1^N X_k\right]}{\partial t} = -\frac{(1-\rho_\infty)[1-\rho(t)]}{(1-\rho_\infty)t}$$
$$\times \sum_{k=1}^N \frac{\partial}{\partial x_k}\left(r_k T(t) \prod_{j=1}^N X_j\right), \quad (18)$$

which, after rearranging and differentiating, becomes

$$\frac{1-\rho_\infty}{1-\rho(t)} \frac{t}{T(t)} \frac{dT}{dt} \prod_{k=1}^N X_k$$
$$= -(1-\rho_\infty) \sum_{k=1}^N \left(r_k \frac{dX_k}{dx_k} \prod_{j\neq k} X_j + \frac{dr_k}{dx_k} \prod_{j=1}^N X_j \right). \quad (19)$$

Dividing through by the term $\prod_{k=1}^N X_k$, we find

$$\frac{dT}{dt} \frac{(1-\rho_\infty)t}{[1-\rho(t)]T(t)}$$
$$= -(1-\rho_\infty) \sum_{k=1}^N \left(r_k \frac{dX_k}{dx_k} X_k^{-1} + \frac{dr_k}{dx_k} \right), \quad (20)$$

which can be separated into an uncoupled system of $N+1$ ordinary differential equations (ODEs) with structures identical to those solved in Eqs. (5),

$$\frac{dT}{dt} = \frac{\lambda}{1-\rho_\infty} \frac{1-\rho(t)}{t} T(t), \quad (21)$$

$$\frac{dX_k}{dx_k} = -\left(\frac{\lambda_k}{(1-\rho_\infty)r_k} + \frac{r_k'}{r_k} \right) X_k, \quad \text{for } k = 1, \ldots, N, \quad (22)$$

where $\lambda = \sum_{k=1}^N \lambda_k$ are the coefficients of separation. The general solution of Eq. (17) is thus

$$f(x_1, \ldots, x_N, t)$$
$$\propto \frac{c}{r} \exp\left[\frac{1}{1-\rho_\infty} \left(\int^t \frac{1-\rho(t')}{t'} dt' - \sum_{k=1}^N \int^{x_k} \frac{dx_k'}{r_k(x_k')} \right) \right], \quad (23)$$

with $c = c_t \prod_k c_k$ and $r = \prod_k r_k(x_k)$.

2. The case of nonzero diffusion

Let Δt be some small time interval and pick $t_0 < t_1$ so that $t_1 - t_0 < \Delta t$ defines some time window of interest. Define $\tau(t) = \frac{(1-\rho_\infty)t}{1-\rho(t)}$. Fixing $\tau_0 = \tau(t_0)$ and $\tau_1 = \tau(t_1)$ and substituting τ_j, $j = 1,2$ for τ in Eq. (16) define solutions $f_0(x_1, \ldots, x_n, t)$ and $f_1(x_1, \ldots, x_n, t)$ whose generators are time independent and whose spatial averages form an envelope for the spatial average of the solution of Eq. (16) as $f_1(t) \leqslant f(t) \leqslant f_0(t)$ for $t \in [t_0, t_1]$. We will solve for the upper envelope solution f_0 explicitly; to solve for f_1 one proceeds identically. Again assuming a product solution of the form $f_0(x_1, \ldots, x_n, t) = T_0(t) \prod_{k=1}^n X_k(x_k)$ and supposing that the diffusion matrix is $\mathbf{D} = \frac{\sigma^2}{2}\mathbf{I} = D\mathbf{I}$, substitution in Eq. (16)

gives

$$\frac{\partial\left[T_0(t)\prod_{i=1}^n X_i\right]}{\partial t} = -\tau_0(1-\rho_\infty)$$
$$\times \sum_{k=1}^N \frac{\partial}{\partial x_k}\left(r_k T_0(t) \prod_{j=1}^N X_j \right)$$
$$+ D \sum_{k=1}^n \frac{\partial^2}{\partial x_k^2}\left(T_0(t) \prod_{j=1}^n X_j \right). \quad (24)$$

Following a similar derivation to Eqs. (18)–(21), one arrives at the system of ODEs,

$$\frac{dT_0}{dt} = \frac{\lambda}{\tau_0} T_0(t), \quad (25)$$

$$\tau_0 D \frac{d^2 X_k}{dx_k^2} + r_k \frac{dX_k}{dx_k} + \left[\frac{\lambda_k}{1-\rho_\infty} + r_k' \right] X_k = 0, \quad (26)$$

where again $\sum_{k=1}^n \lambda_k = \lambda$. The time solution is now given by $T_0(t) = \exp\left(\frac{\lambda}{\tau_0} t\right)$ (recall that this is defined only over $t \in [t_0, t_1]$), while the spatial solutions are much more intricate than those given in Eq. (23). Where $r_k(x_k) = 1$ the solution is given in terms of sines and cosines and it is seen that Eq. (16) simply becomes the heat equation; there is no preferential attachment process here. The case of classical preferential attachment is given by $r_k(x_k) = x_k$, whereupon the spatial equations take the form

$$\frac{d^2 X}{dx^2} - c_1 x \frac{dX}{dx} + c_2 X = 0, \quad (27)$$

where we have set $X = X_k$ for clarity and defined the constants $c_1 = -\frac{1}{\tau_0 D}$ and $c_1 = \frac{1}{\tau_0 D}\left(\frac{\lambda}{1-\rho_\infty} - 1\right)$. This equation is of Hermite type and its solution can be expressed analytically in terms of the confluent hypergeometric function. This does not provide elucidation of the resultant distribution, however; we derive the solution of Eq. (27) in frequency space presently. Defining the Fourier transform by $F(\omega) = F[f](\omega) = \int_{-\infty}^\infty f(x)e^{i\omega x}dx$, transforming Eq. (27) results in the frequency-space differential equation

$$\frac{\partial F}{\partial \omega} = \frac{1}{c_1 \omega}(\omega^2 - c_1 - c_2)F(\omega). \quad (28)$$

The asymptotic solution to Eq. (28) in frequency space is thus (replacing F by F_k)

$$F_k(\omega) \simeq \omega^{-\left(2-\frac{\lambda_k}{1-\rho_\infty}\right)} e^{-\frac{\omega^2}{2\tau_0 D}}. \quad (29)$$

We note the decomposition of $F_k(\omega)$ as a product (in frequency space) of a pure diffusion part and a preferential attachment (power-law) frequency decay; the corresponding time-valued function is a convolution of a diffusion process and the preferential attachment process.

Setting $L^\dagger = -\frac{1-\rho(t)}{t}\mathcal{L}_{\text{div}} + \frac{\sigma^2}{2}\mathcal{L}_{\text{fluc}}$, we can write Eq. (16) as $\frac{d}{dt}f = L^\dagger f$ when the diffusion is uncorrelated (the diffusion matrix is a multiple of the identity). Another way of characterizing solutions to this equation is to solve for a stochastic process that generates the equivalent backward solutions; we search for solutions of $-\frac{d}{dt}f = Lf$. The corresponding partial

differential equation (PDE) is given by

$$-\frac{\partial f}{\partial t} = \frac{1 - \rho(t)}{t}r(x)\frac{\partial f}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}, \qquad (30)$$

defined for $t \in [t_0, T]$ with the final condition $f(x, T) = \phi(x)$. (In dimensions higher than one the extension is clear.) By the Feynman-Kac formula, the solution to this equation is given by

$$f(x, t) = \langle \phi(X_T) | X_t = x \rangle, \qquad (31)$$

where the stochastic process X_t is defined by

$$dX_t = \frac{1 - \rho(t)}{t}r(X_t)dt + \sigma \, dW_t. \qquad (32)$$

Thus analysis of the associated Itô stochastic differential equation (SDE) yields yet another method by which the behavior of the continuum rich-get-richer process can be analyzed.

III. APPLICATION TO MARKET STRUCTURE

A. Firm size

We demonstrate the applicability of our results with a microeconomic analysis of firm revenues. Consider a small time period Δt in which consumers enter a market to purchase an item priced at p. During this time period, we assume each consumer purchases only one item. With probability ρ (likely quite small) a consumer will choose to start their own firm; with probability $1 - \rho$ they will choose to buy the product from an existing firm. Consumers chose a firm from which to buy in proportion to the advertising level of the firm, which is itself proportional to the firm's revenue R. Revenue is given by the equation $R(q) = p(q) \cdot q$, where q is the quantity of the product sold. Firms are price takers, with the market price set at p, so that revenue is $R(q) = pq + c$ for c some constant. Applying the above model, the mean field equation for this process is

$$\frac{\partial f(x, t)}{\partial t} = -\frac{1 - \rho}{t}\frac{\partial}{\partial x}[(px + c)f(x, t)], \qquad (33)$$

where the substitution $q = x$ comes from the restriction that each consumer purchases only one product during the small time interval of study. The solution to this equation is given by

$$f(x, t) \propto t(px + c)^{-1 - \frac{1}{p(1-\rho)}}. \qquad (34)$$

Rewriting this explicitly as $f(R, t) \propto tR^{-1 - \frac{1}{p(1-\rho)}}$ emphasizes the result of a power-law distribution of firms in revenue. This result corresponds with simulation [14,15] and empirical data [13,16]. We note that other attempts to quantify this phenomenon have been either empirical, computational, or statistical in nature; this appears to be a mechanistic model to naturally generate these dynamics.

Figure 3 displays the frequency distributions of U.S. firms with respect to the number of workers from the years 1977 to 2013. That a power law fits these data is well known [13]. As U.S. firms exhibit constant returns to scale [17], this implies a power-law frequency distribution of firms with respect to revenue. We note that the average coefficient β_1 of the log-log fit $\langle \log_{10} f(x) \rangle = \beta_0 + \beta_1 \log_{10} x$ is approximately equal to 1, implying that $\frac{1}{p(1-\rho)} \approx 0$ in Eq. (34). The inset of Fig. 3 displays $\langle f(x, t) \rangle_x$ over the entire date range. These data exhibit linear scaling, as predicted by Eq. (34), with the line
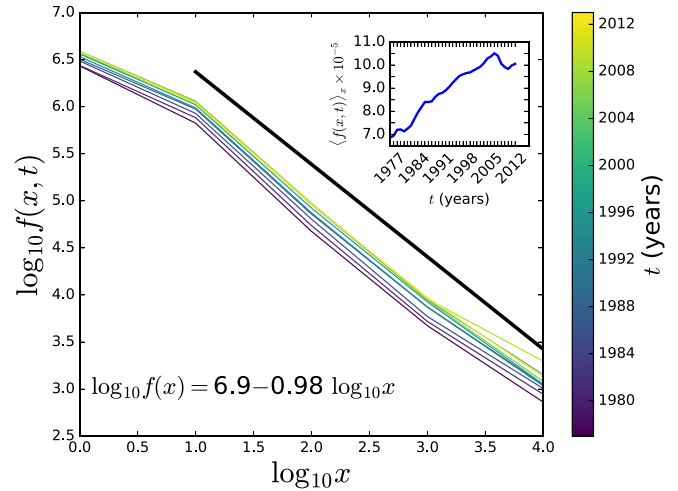

FIG. 3. Power-law size distribution of firms. Distribution of firm sizes by employment from 1977 to 2013, obtained from the U.S. Census Bureau on 28 August 2016 [12]. Note that U.S. firms exhibit constant returns to scale, so this implies a power-law distribution (with identical exponent) in firm income. This result was famously publicized by Axtell in Ref. [13]. The wide binning in the figure is due to the lack of granularity in publicly available U.S. Census data on firms with more than 10^4 employees. The inset displays $\langle f(x, t) \rangle_x$ over the entire date range. These data exhibit linear scaling, as predicted by Eq. (34).

of best fit given by $\langle f(x, t) \rangle_x = -1.98 \times 10^7 + 1.04 \times 10^4 t$ ($R^2 = 0.9204$, $p = 8.042 \times 10^{-21}$). The unpredicted downward trend in $\langle f(x, t) \rangle_x$ in the late 2000's is likely due to unstable conditions in the American economy during this time.

B. Wealth distribution

From Eq. (34), we show that the cumulative wealth distribution of firms exhibits power-law scaling. Defining the wealth kernel to be $w(x, t) = \pi(x, t)f(x, t)$, where $\pi(x, t)$ are the profits resulting from a sale of x items at time t, and imposing a maximum customer base of $x_{\max}$, we have that total system wealth at time t is given by

$$W(t) = \int_{t_{\min}}^{t} \int_0^{x_{\max}} w(x', t')dx' \, dt'. \qquad (35)$$

The functional form of $\pi(x, t)$ is dependent on the functional form of firms' cost function $C(x, t)$. Suppose that firms face identical weakly quadratic costs $C(x, t) = x_0 + C'(x_0)(x - x_0) + \frac{1}{2}C''(x_0)(x - x_0)^2 + O((x - x_0)^3) \sim cx + \varepsilon x^2$, where we assume $0 < \varepsilon \ll 1$. (We choose ε in this range so as to enforce the first-order condition that $\frac{d\pi}{dx} = 0$ has a solution in $\mathbb{R}_+$.) Then, denoting the total wealth of firms with x or more customers at time t by $W_{\geqslant}(x, t)$ and letting $p_{\text{net}} = p - c$, the above equation becomes

$$W_{\geqslant}(x, t) \sim \int_{t_{\min}}^{t} \int_x^{x_{\max}} (p_{\text{net}}x' - \varepsilon x'^2)t'x'^{-1 - \frac{1}{p(1-\rho)}} dx' \, dt'$$

$$\sim p_{\text{net}} \int_{t_{\min}}^{t} t' dt' \int_x^{x_{\max}} x'^{-\frac{1}{p(1-\rho)}} dx'$$

$$\propto p_{\text{net}}t^2 \left(x_{\max}^{1 - \frac{1}{p(1-\rho)}} - x^{1 - \frac{1}{p(1-\rho)}}\right). \qquad (36)$$
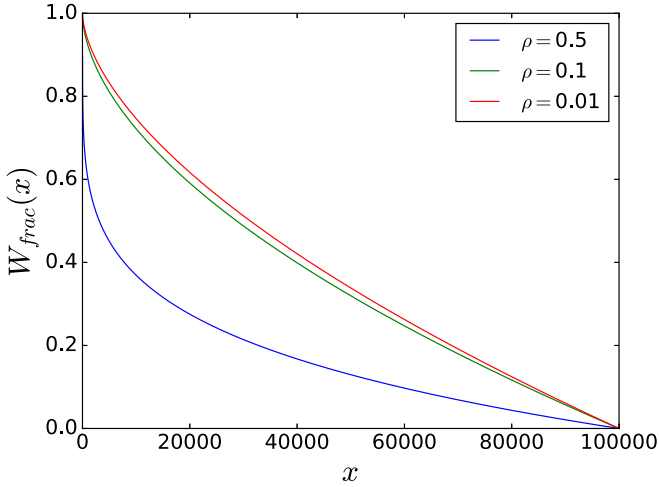
FIG. 4. The fraction of wealth belonging to firms with customer bases greater than or equal to x for $\rho = 0.5, 0.1$, and 0.01. We set the price level $p = 2.5$ for ease of visualization.

Thus the wealth fraction belonging to firms with a customer base greater than or equal to x at time t, denoted $W_{\text{frac}}(x) = W_{\geqslant}(x,t)/W(t)$, displays power-law scaling as a direct result of the preferential attachment process,

$$W_{\text{frac}}(x) \simeq 1 - c_0 x^{1 - \frac{1}{p(1-\rho)}}, \tag{37}$$

where $c_0 \approx (x_{\max}^{1 - \frac{1}{p(1-\rho)}})^{-1}$.

Equation (37) is shown in Fig. 4 for several values of ρ.

IV. CONCLUDING REMARKS AND FURTHER EXTENSIONS

In sum, we have shown how Simon's model, and preferential attachment models more generally, may be extended to the continuum for ease of use as mean field approximations to stochastic processes. We have developed Simon's model in a continuum, expanded upon it by introducing an arbitrary growth kernel $r(x)$ and a time-variant innovation rate $\rho(t)$, and solved the model, discussing the cases in which the general solution satisfies the boundary conditions of the PDE. We are able to find explicit solutions with various growth kernels $r(x)$. Noting that preferential attachment processes may operate in

more than one dimension, we allowed the model to have an arbitrary number of dimensions and solved it there. Finally, we applied the model to the case of consumer accumulation to firms, to demonstrate a theoretical derivation of the power-law distribution of firms by revenue that is observed empirically.

Further extensions to this model could consider the case in which, as treated above, market entrants create their own firms with probability ρ. It might be that market entrants create not a single firm, but multiple firms, or that, in times of economic crisis, firms are removed from the marketplace with probability q. The model could then be described by

$$\frac{\partial f}{\partial t} = -\frac{1-\rho}{t}\frac{\partial}{\partial x}[r(x)f] + q \cdot g[f(x,t),x,t]. \tag{38}$$

Other models could also incorporate past information about the state of the market or a threshold condition via an equation of the form

$$\frac{\partial f}{\partial t} = -\frac{1-\rho}{t}\frac{\partial}{\partial x}[r(x)f]$$
$$+ \int_{x_1}^{x_2} \int_{t_1}^{t_2} h(f(x-x',t-t'),x',t')dt'dx'. \tag{39}$$

As a concluding example of the possible further generalizations, let us consider the problem presented in Eq. (4) in a nonasymptotic setting; consumers form their own firm at a rate $\rho(t)$ with an initial firm intensity given by $\mathcal{I}(x - x_0)$. The model is thus governed by

$$\frac{\partial f}{\partial t} = -\frac{1-\rho(t)}{t}\frac{\partial}{\partial x}[r(x)f] + \rho(t)\mathcal{I}(x - x_0). \tag{40}$$

Again using the method of characteristics, we solve the equations

$$\frac{dx}{dt} = \frac{1-\rho(t)}{t}r(x), \tag{41}$$

$$\frac{df}{dx} = -\frac{r'(x)}{r(x)}f(x,t) + \frac{t\rho(t)}{1-\rho(t)}\mathcal{I}(x - x_0). \tag{42}$$

Solving and again setting the constants of integration A and B to $B = F(A)$, the general solution is given by

$$f(x,t) = \frac{1}{r(x)}\left[F\left(\int \frac{1-\rho(t)}{t}dt - \int \frac{dx}{r(x)}\right)\right.$$
$$\left. + \frac{t\rho(t)}{1-\rho(t)}\int \mathcal{I}(x - x_0)\,dx\right]. \tag{43}$$

[1] H. A. Simon, Biometrika **42**, 425 (1955).

[2] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[3] S. Bornholdt and T. Rohlf, Phys. Rev. Lett. **84**, 6114 (2000).

[4] T. Maillart, D. Sornette, S. Spaeth, and G. von Krogh, Phys. Rev. Lett. **101**, 218701 (2008).

[5] P. S. Dodds, D. R. Dewhurst, F. F. Hazlehurst, C. M. Van Oort, L. Mitchell, A. J. Reagan, J. R. Williams, and C. M. Danforth, Phys. Rev. E **95**, 052301 (2017).

[6] All data and code to recreate figures are available on the lead author's website: https://github.com/daviddewhurst/continuum-preferential-attachment.

[7] J. M. Epstein, in *Handbook of Computational Economics*, edited by L. Tesfatsion and K. L. Judd (Elsevier, Amsterdam, 2006), Vol. 2, Chap. 34, p. 1585.

[8] A. K. Sen, Philos. Public Affairs **6**, 317 (1977).

[9] P. L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).

[10] J. Saldaña, Phys. Rev. E **75**, 027102 (2007).

[11] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).

[12] U.S. firm size data, www2.census.gov/ces/bds/firm/bds_f_sz_release.xlsx (online; accessed 28 August 2016).

[13] R. L. Axtell, Science **293**, 1818 (2001).

[14] R. L. Axtell, in *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems AAMAS '16* (International Foundation for Autonomous Agents and Multiagent Systems, Singapore, Singapore, 2016), pp. 806–816.

[15] L. A. Nunes Amaral, S. V. Buldyrev, S. Havlin, M. A. Salinger, and H. E. Stanley, Phys. Rev. Lett. **80**, 1385 (1998).

[16] J. Zhang, Q. Chen, and Y. Wang, Physica A **388**, 2020 (2009).

[17] S. Basu and J. G. Fernald, J. Political Econ. **105**, 249 (1997).