Behav Ecol Sociobiol (2011) 65:537–546
DOI 10.1007/s00265-010-1102-1

ORIGINAL PAPER

Evolving social influence in large populations

R. Alexander Bentley · Paul Ormerod · Michael Batty

Received: 8 March 2010 / Revised: 28 October 2010 / Accepted: 29 October 2010 / Published online: 1 December 2010

Abstract Darwinian studies of collective human behaviour, which deal fluently with change and are grounded in the details of social influence among individuals, have much to offer "social" models from the physical sciences which have elegant statistical regularities. Although Darwinian evolution is often associated with selection and adaptation, "neutral" models of drift are equally relevant. Building on established neutral models, we present a general, yet highly parsimonious, stochastic model, which generates an entire family of real-world, right-skew socio-economic distributions, including exponential, winner-take-all, power law tails of varying exponents, and power laws across the whole data. The widely used Barabási and Albert (1999) Science 286: 509-512 "B-A" model of preferential attachment is a special case of this general model. In addition, the model produces the continuous turnover observed empirically within these distributions. Previous preferential attachment models have generated specific distributions with turnover using arbitrary add-on rules, but turnover is an inherent feature of our model. The model also replicates an intriguing new relationship, observed across a range of empirical studies, between the power law exponent and the proportion of data represented in the distribution.

Keywords Neutral theory · Human dynamics · Scaling · Pop music · Markets · Culture evolution · Baby names · Cultural transmission · Power laws · Fashion · Random copying

Communicated by Guest Editor A. Houston

This contribution is part of the Special Issue "Mathematical Models in Ecology and Evolution: Darwin 200" (see Marshall et al. 2010).

R. A. Bentley (✉) · P. Ormerod
Department of Anthropology, Durham University,
Dawson Building, South Road Durham,
Durham DH1 3LE, UK
e-mail: r.a.bentley@durham.ac.uk

P. Ormerod
Volterra Consulting,
135c Sheen Lane,
London SW14 8AE, UK
e-mail: p.ormerod@volterra.co.uk

M. Batty
Centre for Advanced Spatial Analysis,
University College London,
1–19 Torrington Place,
London WC1E 6BT, UK
e-mail: m.batty@ucl.ac.uk

Introduction

In behavioural ecology, the evolution of cultural traditions is often viewed as a process of functional adaptation, achieved at the level of individual selection, through the currency of individual costs and benefits (cf. Weissing et al. 2010). Human behavioural ecology is a tradition in itself, of exploring social evolution in terms of energy budgets and the adaptive value of behaviour within a particular physical environment (Steward 1955; Winterhalder and Smith 2000; Cronk and Gerkey 2007; Houston 2010). In this tradition, there is often an assumption that the behavioural variation needed to "solve" an adaptive problem available through behavioural plasticity gives humans a range of possible strategies from which they can rationally choose and then apply to particular situations in the attempt to optimise energy expenditure, time expenditure, reproductive or other benefits (Boone and Smith 1998; Winterhalder and Smith 2000; Holden et al. 2003; Borgerhoff Mulder et al. 2009; Houston 2010). Hunting or gathering optimises the acquisition protein and/or calories, while preferences for human attractiveness reflect attributes of health or reproductive fitness, such as an optimal waist-to-hip ratio in a woman's figure, or symmetry in a man's face.

Human behavioural ecology can therefore predict broadly the functional constraints on human decisions, but within those constraints, there are often many equivalent choices, particularly in "fashionable" activities like pottery-making, music, art, humour, language and material culture. Cultural evolution acceptable variation is present even among such adaptively crucial activities as mating and procuring food. The "ideal" female figure that appeals to US college students looks like someone who is "skinny in the waist and has diarrhoea" to a Yomybato man of the Peruvian rain forest (Yu and Sheperd 1998). Even group norms of economic fairness (i.e. results in the Ultimatum game) vary substantially cross-culturally (Henrich et al. 2005).

Compared to the small-scale societies of human prehistory, the range of choice in modern Western consumer culture is many orders of magnitude larger (Beinhocker 2006). There can be so many nearly equivalent choices (think of all the similar brands of mobile phones, shoes, and soap powders), that there is almost nothing to distinguish them inherently. Given too much choice, even physical attractiveness and mating strategies (cf. Hasson and Stone 2010) are affected—online social networking, "speed dating" and urban population sizes provide orders of magnitude more choice than in a small prehistoric hunter-gatherer society, where choice of mate was relatively minimal (if not effectively pre-arranged through kinship ties).

Within a range of functional constraints, then, human choice is much more free to be determined through social value, especially as humans have evolved into social animals *par excellence* (Dunbar and Shultz 2007; cf. Marshall 2010). Few behavioural choices can be explained exclusively in terms of individual costs and benefits involving calories, or energy, time or other currency. Even hunting, which is not the most reliable means of procuring protein—females can often supply it more reliably by gathering nuts, for example (Lee and DeVore 1976)—is simultaneously a crucial means for males to socialise and acquire prestige (Wiessner and Schiefenhövel 1997; Henrich and Gil-White 2001).

Social learning is a complex process, as there is a wide range of strategies people use to choose what or whom to copy (Laland 2004; Henrich and Gil-White 2001), which facilitates the development of different norms and beliefs between groups (Boyd et al. 2010). In small-scale societies where a detailed ethnography can be done, these different strategies or biases might be delineated. Among large populations, however, like modern mass media markets, or populations of Internet users, for example, virtually *all* of these copying biases will exist, in different proportions. Just about everyone will have a particular, personal reason for choosing a first name for their child, popular book, or even ideological belief (in modern developed countries where choice is more individual). Also, the vast amount of specialised knowledge in the world is now perhaps a billion times (give or take an order of magnitude) the technological variation of a prehistoric hunter-gatherer community (Beinhocker 2006).

At the scale of large populations, it is virtually impossible to track individual biases from aggregated datasets such as a census data, popularity statistics, sales figures, and so on. The new scale is, in effect, a new phenomenon (Anderson 1972). Just as chemists would use the Ideal Gas Law to predict the pressure of a gas rather than try to tally the physical energies of all the individual molecules, social scientists looking at populations of social learners need to abandon hope of tracking individual copying biases and perform analyses in a more explicitly statistical way. We thus focus on populations rather than individuals. Complex cultural traditions, as Boyd and Richerson (2005: 16) point out, "are the product of a population of minds … In the absence of such a population, the costly structures necessary for accurate imitation are useless".

For large populations where individual learning biases are virtually intractable, we need an appropriate approach—a bit like the Ideal Gas Law—that treats the population *as if* all the possible biases and individual rationales balance each other out. This is especially appropriate where there is a large range of virtually equivalent choices in terms of functional value—as in spoken or sign language (Hauser et al. 2002), prehistoric pottery decorations (Neiman 1995; Shennan and Wilkinson 2001), or writing, for example, as well as in modern fashionable names and words (Hahn and Bentley 2003; Berger and Le Mens 2009), leisure activities (Bentley and Ormerod 2009), and music preferences in a social environment (Salganik et al. 2006).

For behaviours that are socially learned, cultural drift is becoming increasingly evident (Koerper and Stickel 1980) to the point where population size and social learning can become one of the determinants of culture itself (e.g., Renfrew and Scarre 1998; Shennan 2000; Powell et al. 2009). One of the outcomes of social learning at the scale of larger populations is that cultural element frequencies are characterised by stochastic change, rather than by usefulness and adaptation. This was recently demonstrated, for example, by an Internet-based experiment on music downloading (Salganik et al. 2006), and studies of baby names (Hahn and Bentley 2003; Berger and Le Mens 2009).

This is why a simple model of random copying among individuals (with occasional innovation), also known as the "Neutral model", can fit many of the data patterns of cultural change, among both humans and social animals such as birds (Neiman 1995; Shennan and Wilkinson 2001; Lachlan and Slater 2003; Byers et al. 2010). Although neutral traits appear to be fundamentally unpredictable in terms of the frequencies of specific variants, they are collectively characterised by a number of statistical regularities involving the variation within the set of adopted variants (Neiman 1995; Lipo et al. 1997; Shennan and Wilkinson 2001), regular turnover among the most popularity variants (Bentley et al. 2007), and in long-tailed distributions of the variant frequencies (Bentley et al. 2004; Mesoudi and Lycett 2009).

This model is equivalent to an evolutionary neutral model (cf. Hahn and Bentley 2003), except with the added memory parameter. In our model, each individual holds one "idea", and that idea may be copied by another individual. Hence the ideas (the parallels of phenotypes in evolutionary neutral theories) replicate with probability proportional to the number of individuals having that idea, although importantly, the actual replication varies from that frequency-dependent probability through random chance.

The Neutral model has thus shown great promise for explaining long-tailed distributions in socio-cultural data, but so far it has not quite been able to replicate the wide range of long-tailed distributions that exist in society. Recently, Mesoudi and Lycett (2009) explored the effect of incrementally adding a conformity bias to the Neutral model, which essentially skewed the variant frequency distribution towards what is sometimes known as the "winner-take-all" distribution, which Bentley and Shennan (2003) had expected under frequency-dependent copying. This modification was an advance, but there is still a much broader range of long-tailed distributions left to explain (Newman 2005; Clauset et al. 2009).

We thus propose a modified Neutral model, based upon individual agents who are boundedly rational and are influenced by the behaviour of other agents in terms of their decision-making. In other words, the agents act with social purpose, fundamentally different from physical or biological phenomena where the agents (or particles) are incapable of intent. The key improvement that we make is to introduce a "memory parameter", which allows agents to look at *any* number of previous decisions by other agents, from just the previous period alone (at one extreme) to all previous periods (at the other extreme). We return to this point below.

The model provides four advances on previous models:

- It can generate a wide range of right-skew distributions observed in cultural, economic and social situations from different combinations of its two parameters.
- The widely used Barabási-Albert (B-A) model (Barabási and Albert 1999) of preferential attachment is a special case of this general model.
- In terms of power law fits, there are two essential statistics, the exponent a and the fraction f of the total observations over which the power law is believed to hold. The model can replicate both observed exponents a and the fraction f from real-world observations (Newman 2005; Clauset et al. 2009).
- Many real-world right-skew distributions exhibit constant turnover in the rankings of their constituents even if their functional form is time-invariant (Batty 2006; Bentley et al. 2007). Our model is capable of generating such turnover without recourse to self-fulfilling rules such as "aging" or variable "fitness" of the individual elements (Newman 2005).

We stress that our model is not a network model, and therefore *not* a modified B-A model as such. Whenever a network model is applied to social dynamics, two analogies are possible:

- The nodes represent the agents (e.g., academic papers), and the links are the relationships between them (citations)
- The links represent agents (e.g., consumers), and the nodes represent the choices the agents make (e.g., purchased brands)

Whatever the representation, new nodes in B-A models link with probability $p(k)$ to existing nodes, where $p(k)$ is a function of the degree k of the existing node (and possibly time t as well, in aging models). Agents and nodes are inseparable, connected by the network. Our evolutionary model is different because the agents and their ideas are *separable*: each new agent adopts its idea either by copying another agent, or by inventing an idea of its own.

Methods

The model builds on previous versions of the cultural Neutral model (e.g. Neiman 1995; Shennan and Wilkinson 2001; Hahn and Bentley 2003; Bentley et al. 2004, 2007; Mesoudi and Lycett 2009), but with several key modifications, especially the variable memory parameter mentioned above.

Consider a model populated initially by N agents located in some abstract space such as a sequence of index numbers. Depending on the phenomenon, each location is an abstract representation; it could refer to the city where a firm chooses to locate itself, but it could equally well refer to the product a consumer chooses, or the idea or fashion that a person follows. We define the *size* of a location as the number of agents at that location.

The model proceeds in a series of steps. In each step, n new agents enter the model, where the number n is fixed as a parameter in each solution of the model. With probability 1-μ, an agent copies the choice of location from that of an existing agent within the previous m time steps, or else with probability μ, the agent innovates by choosing a unique new location at random. In other words, the agent either copies an existing agent from the last m steps, or chooses a new location. As described above, the choice for copying is made from the pool of agents, such that the probability a location is copied is proportional to the number of agents already located there.

Here we restrict our exploration to two key parameters of the model, m and μ, by choosing convenient values for N and n. The "memory" parameter m determines the number of steps of the previous decisions of other agents over which an agent looks when making its decision. The "innovation" parameter μ determines the fraction of the agents who decide to take a completely new decision rather than replicating one of the decisions made by other agents.

Results

The Neutral model as previously studied (e.g., Bentley et al. 2007; Mesoudi and Lycett 2009) is actually a specific version of this new, more general model, with m=1 (i.e., memory only of the immediately preceding step). For the special case of n=N and m=1, analytical solutions demonstrate a power law distribution (Evans 2007) for $N\mu$ equal to or slightly greater than 1. For m=1 and $N\mu$<<1, this gradually converges on a winner-take-all distribution as $N\mu$ approaches zero.

The case where m=all is a further special category of the model, where extinction or obsolescence does not occur. In this case, we can achieve different power law slopes by varying n and μ. Figure 1 shows, for example, that we can match the B-A preferential attachment model (Barabási and Albert 1999), obtaining a power law exponent a~3 over the entire distribution, by using m=all with N=1, n=10, t=20,000, and μ=0.6.

For socio-cultural phenomena, however, we expect memory to be limited, and thus m in general to take values below the special case of "all". So while we define the model to allow m to take any value between 1 and all, we explore here a limited range, from m=1 to m=100 time steps of limited memory. The combined effect of varying m along with varying the innovation parameter μ generates both a wide range of right-skew distributional forms and turnover of rankings of locations within those distributions. Considerable anthropological and socio-economic evidence exists (e.g., Eerkens 2000; Diederen et al. 2003; Srinivasan and Mason 1986; Larsen 1961; Rogers 1962) on the plausible values for μ being no greater than 0.1.

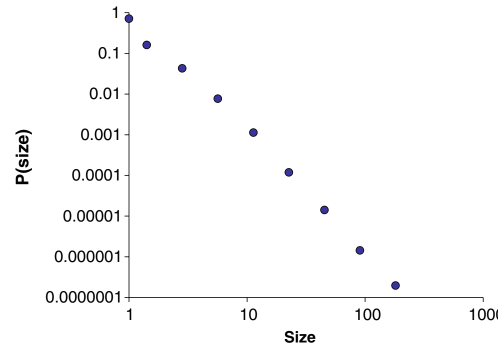

Fig. 1 The power law generated from the preferential attachment version of the model. The probability distribution for a typical model run uses N=1, n=10, t=20,000, μ=0.6, and m=all (where the generated sizes are logarithmically binned). The exponent a for the power law is −2.9 (R^2=0.996), matching that reported (also by ordinary least-squares regression) for preferential attachment models (Barabási and Albert 1999)

Figure 2a plots typical solutions of the model using acceptable values of μ, while varying m (holding N=1,000, n=100 and showing the results at time step 1,000). Aside from the selected results shown in this figure, the model produces additional results ranging from a winner-take-all outcome, to a power law over the entire distribution (exponent a~1.5) to a power law fitted to the tail of varying exponent. Figure 2b illustrates how the model parameters can be selected so that the results match real-world right-skew distributions, such as religions, website subscriptions, word use, names, and author citations.

Regularity in the long tail

Table 1 lists power law tail exponents a for various recently collated social data sets (Newman 2005; Clauset et al. 2009) along with the fraction f (=n_{tail}/n) of total observations in the tail, with the tail defined as the n_{tail} rightmost datum points on the rank-size distribution. A striking, and previously unreported, feature of these estimates is the relationship between a and f, where these data reveal a clear inverse correlation. The smaller the fraction f of the distribution best-fit to a power law tail, the larger the exponent a of that tail. The least-squares fit is a=1.54$f^{-0.156}$ (R^2=0.952, omitting the one outlier of email address book sizes because as many such lists are generated by automated computer algorithms, we consider this to be an unreliable observation).

Figure 3 plots this relationship in the empirical data along with the least-squares fit using the model, as solved 100 times, for each of μ=0.05, 0.06 and 0.07, with m=30 in each case (and N=1,000, n=100, t=1,000). The results

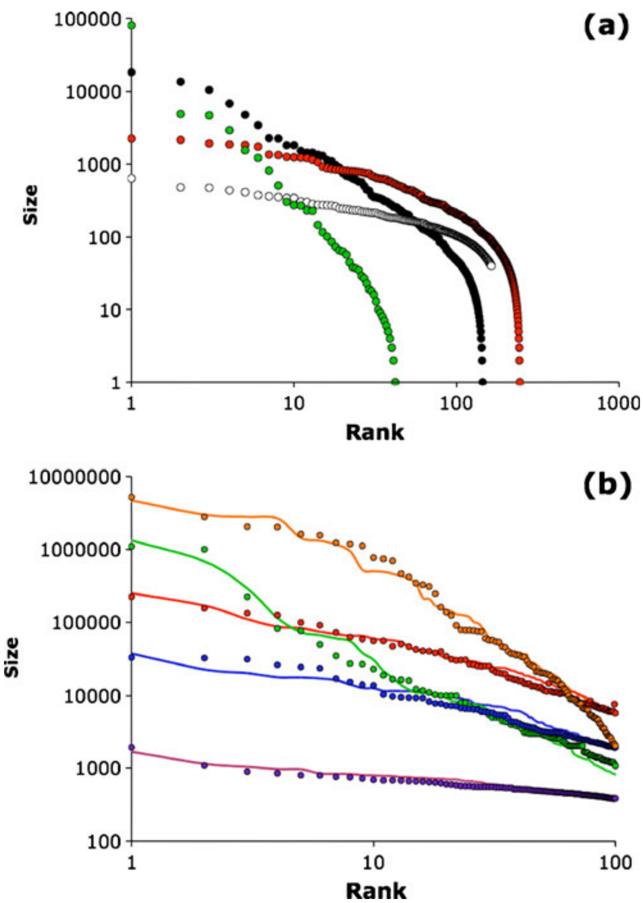

Fig. 2 Log-log plots of rank and size. **a** for typical model solutions with N=1,000, n=100, t=1,000 and: μ=0.01, m=1 (*black*); μ=0.01, m=100 (*red*); μ=0.08, m=100 (*white*); μ=0.0001, m=2 (*green*). **b** For real-world top 100 ranked lists (*dots*) versus model results (*lines*). Top 100 lists include: male baby name frequency (per million) in the 1990 US census (*blue*), RSS feed subscriptions 2001–2008 (*orange*), English words (*red*), cited economists 1993–2003 (*purple*), and religions in thousands of adherents (*green*). With N=1,000, the model fits were made with μ=0.001, m=50, n=200, t=4,000 for names, μ=0.00002, m=6, n=2,500, t=10,000 for RSS feeds, μ=0.00025, m=85, n=100, t=1,100 for cited economists, μ=0.004, m=4, n=450, t=8,000 for words, and μ=0.0007, m=2, n=100, t=4,000 for religions. Data sources: Male baby names from www.census.gov, cited economists from www.in-cites.com, religious adherents from www.adherents.com, RSS feeds from radio.xmlstoragesystem.com/rcsPublic, and English words from www.bckelk.ukfsn.org/words/uk1000n.html

show $a=1.56f^{-0.155}$ (R^2=0.975), very similar to the data-based relationship.

The distribution of turnover

The model also produces continual turnover through time for any given distribution, as demonstrated by the distributions of life-spans within ranked lists (life-span being the number of time steps a location spends on the list) as in Fig. 4a. This resembles the life-spans of real-world social and economic fat-tail distributions in Fig. 4b. The memory parameter m again expands the power of the model. Although turnover has already been demonstrated (Evans 2007) for the special case m=1, different values of m are needed to account for empirically observed turnover.

Figure 5 shows distributions generated of life-spans in the top 5 (i.e., number of time steps location spent in the top 5 most popular) in tests with increasing memory m and invention rate μ (the plots in Fig. 5a-d show progressively larger values of m, while each plot shows five different orders of magnitudes for μ). The effect of increasing memory m is to reduce the effect of innovation μ such that when m= all (Fig. 5d), the distributions are all virtually the same form (albeit with different finite limits to the cumulative number of locations).

Discussion

Socially learned behaviour that could be called "cultural" is observed in animals as diverse as fish (Laland 2004), birds (Lynch and Baker 1994; Slater and Ince 1979; Lachlan and Slater 2003), and of course primates and other species (Galef and Laland 2005). What sets humans apart is that humans are more accurate and complex social imitators than any other animal species (Boyd and Richerson 2005). Social influence is arguably ubiquitous among the human species (Dunbar and Shultz 2007). In fact, rather than the agent's cost-benefit analysis that has served as a null hypothesis for rationality

Table 1 Power-law fits determined by Clauset et al. (2009) among socio-cultural data sets. Parameters include the number of observations n, the maximum observed value x_{max}, the number of observations in the tail n_{tail} and the minimum value in the tail x_{min}

Quantity	n	x_{max}	x_{min}	a (alpha)	n_{tail}	$f=n_{tail}/n$
Intensity of wars	115	382	2.1±3.5	1.7±0.2	70±14	0.609
Religious followers (×10^6)	103	1,050	3.85±1.60	1.8±0.1	39±26	0.379
Word count	18,855	14,086	7±2	1.95±0.02	2,958±987	0.157
City population (×10^3)	19,447	8,009	52.5±11.9	2.37±0.08	580±177	0.030
Terrorist attack severity	9,101	2,749	12±4	2.4±0.2	547±1,663	0.060
Surname frequency (×10^3)	2,753	2,502	112±41	2.5±0.2	239±215	0.087
Paper citations	415,229	8,904	160±35	3.16±0.06	3,455±1,859	0.008
Email address books	4,581	333	57±21	3.5±0.6	196±449	0.043
Papers authored	401,455	1,416	133±13	4.3±0.1	988±377	0.002

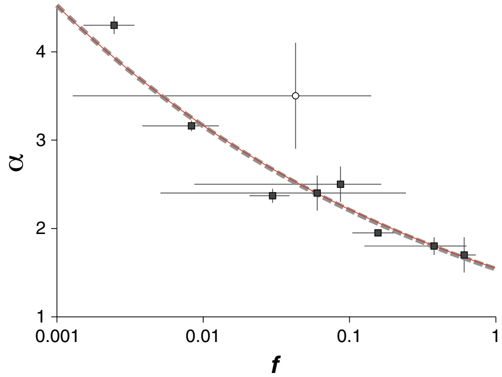

Fig. 3 The power law tail exponent *a* versus the fraction *f* of total observations represented by the tail. The dots show power law tails calculated for various real-world socio-cultural data sets (see Table 1 for values and errors), whose relationship (*dashed grey curve*) can be approximated by $a=1.54f^{-0.156}$ ($R^2=0.952$ except for the outlier—the *open circle*—from email lists). The thin red curve shows the least-squares fit from 300 runs of our theoretical model which gives $a=1.56/f^{-0.155}$ ($R^2=0.975$). Exponents have been estimated using maximum likelihood (Newman 2005; Clauset et al. 2009)

for over a century, an alternative is that each agent uses (consciously or not) the decisions of others as a basis for his or her own decisions.

Applying this to social animals, Laland (2004) presents a useful overview of social-learning strategies, including copying the majority, copying kin, friends, or older individuals, and copying behaviours that are rare, or successful, or better. In humans, this would include more complex social cues, such as copying people of high prestige or status (Henrich and Gil-White 2001). In trying to understand emergent phenomena in modern popular culture, however, it would be virtually impossible for us to trace all these different biases at the population scale. For this reason, we have presented here a model with no biases at all, that is one where the copying process was completely neutral. We can add biases, as was done for a simpler version of the neutral model with fewer parameters (Mesoudi and Lycett 2009), but just by adding the memory parameter as we have done here, we find that our modified neutral model has new complexity that we need to understand before exploring the effects of additional parameters representing biases.

The social-influence model we have presented allows choices among multiple possible alternatives, which rise and fall in relative popularity over time, rather than binary, "either–or" decisions (cf. Kacelnik et al. 2010). This is truly reflective of human interactions such as the choice of a popular name for a child, the citation of an academic paper, or movement to a city where others have chosen to live. Indeed, these phenomena are inherently defined by the past decisions of others, without which there would be no cities, familiar names, or popular culture.

The model also offers an alternative to the many modified Barabási-Albert (B-A) models of the last 10 years for the right-skew nature of income, words, scientific papers, and city sizes, for example (Newman 2005; Clauset et al. 2009; Simon 1955a; Zipf 1949; Price 1965; Pareto 1907). In the statistical sciences, particularly statistical physics, the recent explosion of interest in such distributions for social phenomena includes internet links (Huberman and Adamic 1999; Barabási and Albert 1999), author citations (Redner 1998), sexual partners (Liljeros et al. 2001), and firm sizes and their extinctions (Axtell 2001; Ormerod 2006) amongst many others.

With socio-economic phenomena, the detailed debate over the exact form of these distributions—for example,

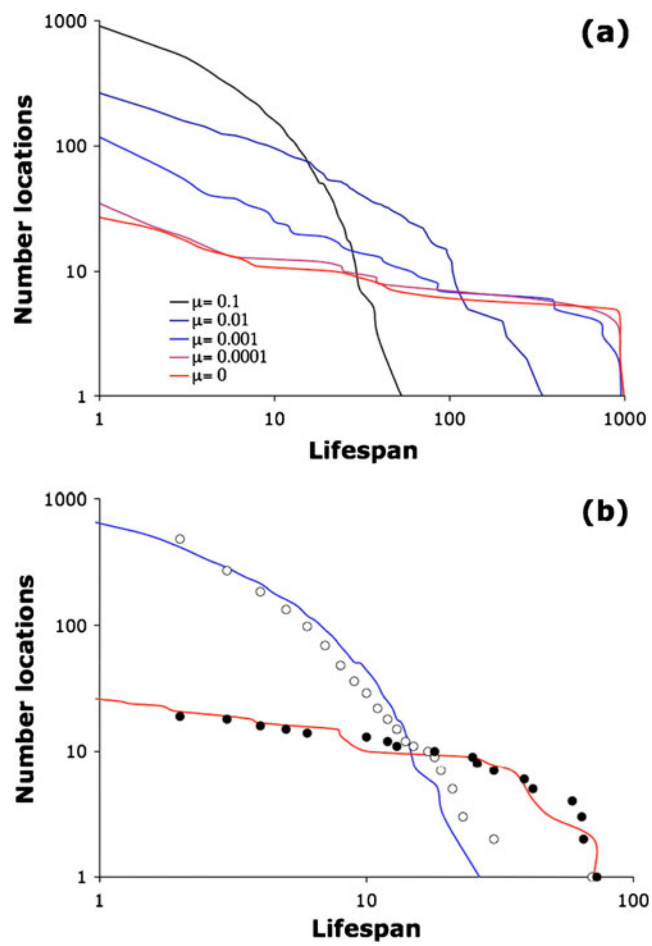

Fig. 4 Life-spans of individual locations. **a** Typical model runs, showing the cumulative distribution of number of time steps spent in the top 5 for model runs of 1,000 time steps with $N=1{,}000$, $n=100$, and $m=1$. **b** Life-spans of UK Number One Hits (www.theofficialcharts.com) for 1956–2007 (*open circles*), versus the model, $m=1$, $\mu=0.1$ (*blue line*), and *t* years in the Top 5 US boys' names (www.ssa.gov/OACT/babynames), 1907–2006 (*filled circles*) versus the model, $m=10$, $\mu=0.001$ (*red line*). Since the temporal units are arbitrary, the modelled life-spans were divided by two to match the albums, and divided by ten to match the names (which on the log-log plot slides the distribution to the left)

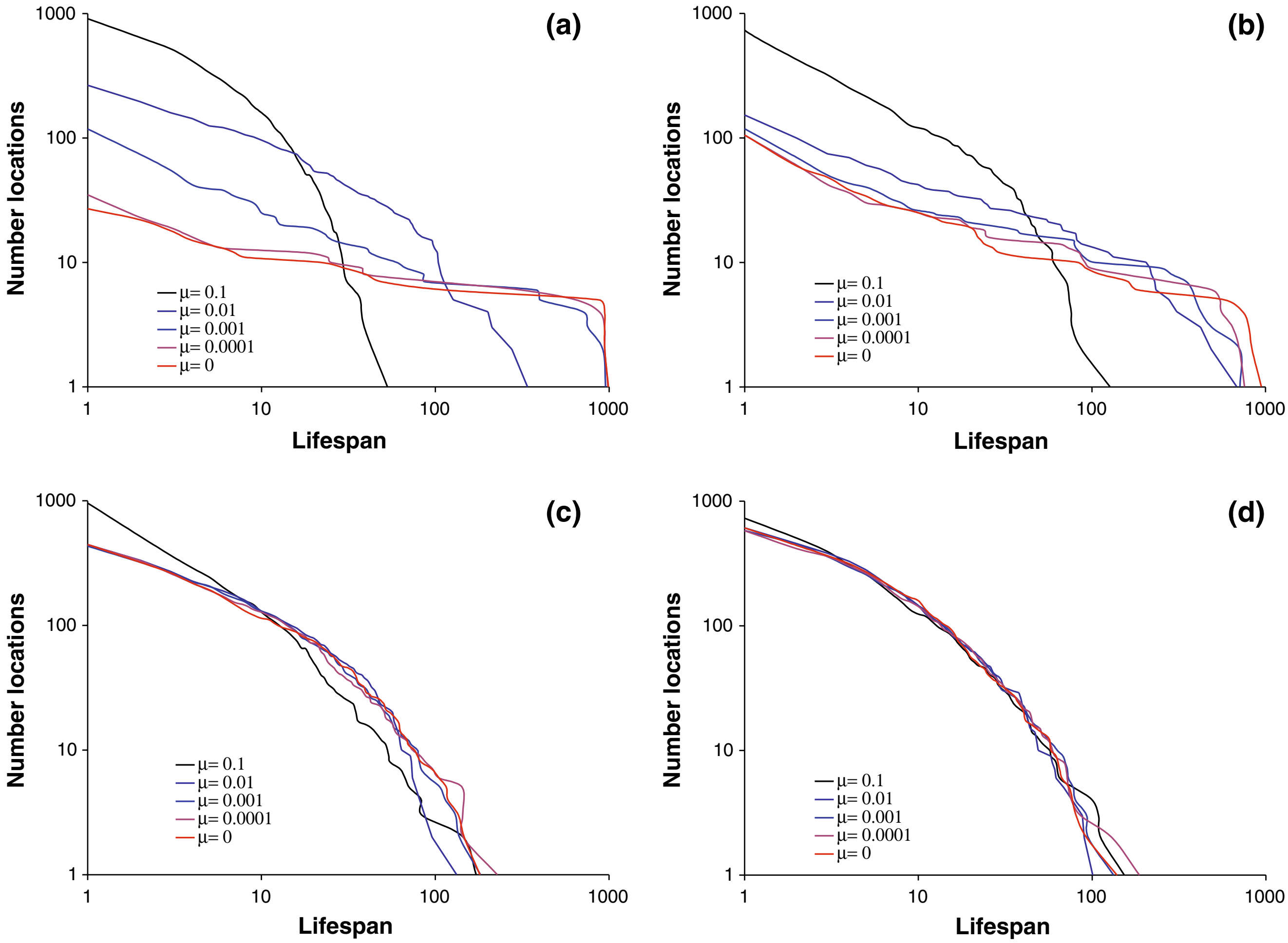

Fig. 5 Modelled life-spans of individual locations. This shows the cumulative distribution of number of time steps spent in the top 5 for model runs of 1,000 time steps with N=1,000, n=100 for **a** m=1, **b** m=10, **c** m=100, **d** m=*all*. The values of μ shown are 0.1 (*black*), 0.01 (*dark blue*), 0.001 (*light blue*), 0.0001 (*purple*) and 0 (*red*)

power laws versus similar fat-tailed functions such as the stretched exponential (Newman 2005; Clauset et al. 2009; Laherrère and Sornette 1998; Perline 2005)—often involves the characterisation of the distribution at a point in time. This neglects the importance of dynamics and the underlying behaviour (Borgatti et al. 2009; Batty 2006), which gives rise to changes over time within any given distribution. Simon (1955a) argued that right-skew distributions were so widespread that their key similarity was likely to be "in the underlying probability mechanisms" that led to their generation. This is clearly the case but, as noted in the social sciences for over a century (Borgatti et al. 2009), it is inherently a description of macro phenomena, without an explanation for the individual behaviour that gives rise to emergent properties.

Within such distributions, there is constant change. In other words, the distributions are not merely long tailed, but dynamic as well. Whilst large numbers of papers have been written on the right-skew nature of the distributions, the turnover within these distributions remains poorly understood. This is an issue on which researchers in evolution and human behaviour can make great advances. Econophysics models struggle with change, and yet change is the essence of evolutionary models. Classical consumer choice theory, involving rational agents making independent cost-benefit decisions, often does not work for populations in which people influence each other's choices (e.g., Ormerod 2006; Beinhocker 2006).

Our model derives from evolutionary neutral models originating over 40 years ago which, ironically, may apply better to selected cases of cultural evolution than to biological molecular evolution (see Hahn 2008). Although for most non-human organisms, behavioural evolution is often associated with natural selection, sexual selection or kin selection (e.g., Edwards 2010; Hasson and Stone 2010; Marshall 2010; Ratnieks et al. 2010; Weissing et al. 2010), our application of an evolutionary neutral model towards modern human fashions achieves several key effects, including matching a range of real-world popularity distributions, their turnover through time, and the empirical match to the a-f relationship in Fig. 3. These real-world effects have never been matched by a single, simple model. The model we have presented can generate not only a wide range of long-tailed distributions but a constant turnover of the constituent agents within any given overall rank-size distribution. It is also able to replicate a newly identified empirical relationship whereby the power law exponent increases as the proportion of data in the tail falls.

The model is quite general, despite using only two main parameters. Varying the parameter values can yield a range of distributions, such as a power law over the whole sample, a power law only in the tail, and a winner-take-all outcome. This combination of results makes this model unique among the many alternatives that can produce power laws. The most commonly proposed processes such as preferential attachment, proportionate effect based on Gibrat's principle, the "Matthew effect" and the Yule process (Newman 2005; Clauset et al. 2009; Batty 2006; Yule 1925; Simon 1955b), produce power laws from the positive feedback introduced by interactions between individual agents. But these "rich get richer" models have not been able to account for flux in the constituents of the ranked distribution (Batty 2008), either when growth is one of strict preferential attachment or even when growth is proportionate to a stochastic rate independent of size (Gibrat 1931).

In related network approaches, such as B-A models with aging (e.g., Dorogovtsev and Mendes 2000), the probability of the choice itself (at the node) diminishes with its age. This is less appropriate for our own interest in the adoption of ideas, where an idea does not go extinct because it is old, but because no one uses it anymore. In many cases the opposite is true; the oldest English words, for example, are actually the ones most frequently spoken today (Pagel et al. 2007). It is more appropriate and effective to model the limited memory of the choosers, as our model does, rather than the aging of the choices. A recent network model (Hajra and Sen 2006) does, in fact, briefly explore limited memory, but again equating nodes with ideas may work for the specific application to authors and their citations of other authors (Gallegati et al. 2006), but does not work at all for the much more general cases that we have explored here, where the ideas and their choosers are entirely separable. By contrast, in our evolutionary model the agents and their ideas are suitably separable: each new agent adopts its idea either by copying another agent, or by inventing an idea of its own.

Social scientists have been critical of modelling social and economic data by mapping onto known phenomena in physics without considering realistic behavioural motivations of the agents (Borgatti et al. 2009; Gallegati et al. 2006; Bentley and Shennan 2005). As a step in this direction, our model captures two fundamental motivations, the imitation of others and novelty in invention.

Compared to simpler but less flexible versions of this model (Hahn and Bentley 2003; Bentley et al. 2007; Byers et al. 2010), a crucial new variable appears to be the memory m, which reflects different time frames to which agents will refer in different contexts. In terms of pure fashion markets such as popular music for example agents take into account only the most recent decisions of others and hardly ever those of several months or even weeks ago. However in choosing where to locate geographically, for example, a firm or a person in a city will implicitly be using information from many previous time steps with respect to the decisions made by others.

Generating a range of long-tailed distributions with dynamic turnover, these features distinguish this model from the standard socio-economic science model of individual rational behaviour where social influence is the exception to the rule (as in, for example, "irrational" stock market bubbles or real estate crises). With its unrealistic psychological assumptions (Kahneman 2003) and inconsistencies with experimental results (Smith 2003), the standard model suffers from a neglect of social influence, even in its modern form which permits, for example, asymmetry in the amount of information possessed by different agents (Akerlof 1970; Stiglitz 2002), the cost of gathering information (Stigler 1961), and imperfections in gathering and processing information (Simon 1955a).

Acknowledgments We thank the Leverhulme Trust for the "Tipping points" programmatic bid. Batty was partially supported by the EPSRC Spatially Embedded Complex Systems Engineering Consortium (EP/C513703/1). Amy Heineike, formerly of Volterra and now at George Mason University, provided programming support.

Conflict of Interests The authors of this manuscript declare no conflict of interests

References

Akerlof GA (1970) The market for 'lemons': quality uncertainty and the market mechanism. Q J Econ 84:488–500
Anderson PW (1972) More is different. Science 177:393–396
Axtell RL (2001) Zipf distribution of U.S. firm sizes. Science 293:1818–1820
Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286:509–512
Batty M (2006) Rank clocks. Nature 444:592–596
Batty M (2008) The size, scale, and shape of cities. Science 319:769–771
Beinhocker E (2006) The origin of wealth. Harvard Business School, Cambridge
Bentley RA, Ormerod P (2009) Tradition and fashion in consumer choice. Scot J Polit Econ 56:371–381
Bentley RA, Shennan SJ (2003) Cultural evolution and stochastic network growth. Am Antiquity 68:459–485

Bentley RA, Shennan SJ (2005) Random copying and cultural evolution. Science 309:877–879

Bentley RA, Hahn MW, Shennan SJ (2004) Random drift and culture change. Proc R Soc B 271:1443–1450

Bentley RA, Lipo CP, Herzog HA, Hahn MW (2007) Regular rates of popular culture change reflect random copying. Evol Hum Behav 28:151–158

Berger J, Le Mens G (2009) How adoption speed affects the abandonment of cultural tastes. Proc Natl Acad Sci USA 106:8146–8150

Boone JL, Smith EA (1998) Is it evolution yet? A critique of evolutionary archaeology. Curr Anthropol 39:S141–S173

Borgatti SP, Mehra A, Brass DJ (2009) Network analysis in the social sciences. Science 323:892–895

Borgerhoff Mulder M, Bowles S, Hertz T, Bell A, Beise J, Clark G, Fazzio I, Gurven M, Hill K, Hooper PL, Irons W, Kaplan H, Leonetti D, Low B, Marlowe F, McElreath R, Naidu S, Nolin D, Piraino P, Quinlan R, Schniter E, Sear R, Shenk M, Smith EA, von Rueden C, Wiessner P (2009) Intergenerational wealth transmission and the dynamics of inequality in small-scale societies. Science 326:682–688

Boyd R, Richerson PJ (2005) The origin and evolution of culture. Oxford University, New York

Boyd R, Richerson PJ, Henrich J (2010) Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. Behav Ecol Sociobiol. doi:10.1007/s00265-010-1100-3

Byers BE, Belinsky KL, Bentley RA (2010) Independent cultural evolution of two song traditions in the chestnut-sided warbler. Am Nat 176:476–489

Clauset A, Shalizi CR, Newman MEJ (2009) Power-law distributions in empirical data. SIAM Rev 51:661–703

Cronk L, Gerkey A (2007) Kinship and descent. In: Dunbar RIM, Barrett L (eds) The Oxford handbook of evolutionary psychology. Oxford University Press, London, pp 463–478

Diederen P, van Meijl H, Wolters A, Bijak K (2003) Innovation adoption in agriculture. Cah Économie Sociol Rurales 67:30–50

Dorogovtsev SN, Mendes JFF (2000) Evolution of networks with aging of sites. Phys Rev E 62:1842

Dunbar RIM, Shultz S (2007) Evolution in the social brain. Science 317:1344–1347

Edwards AWF (2010) Mathematizing Darwin. Behav Ecol Sociobiol. doi:10.1007/s00265-010-1122-x

Eerkens JW (2000) Practice makes within 5% of perfect. Curr Anthropol 41:663–668

Evans TS (2007) Exact solutions for network rewiring models. Eur Phys J B 56:65–69

Galef BG, Laland KN (2005) Social learning in animals: empirical studies and theoretical models. BioScience 55:489–499

Gallegati M, Keen S, Lux T, Ormerod P (2006) Worrying trends in econophysics. Phys A 370:1–6

Gibrat R (1931) Les inégalités économiques. Librarie du Recueil Sirey, Paris

Hahn MW (2008) Toward a selection theory of molecular evolution. Evolution 62:255–265

Hahn MW, Bentley RA (2003) Drift as a mechanism for cultural change: an example from baby names. Proc R Soc B 270:S120–S123

Hajra B, Sen P (2006) Modelling aging characteristics in citation networks. Phys A 368:575–582

Hasson O, Stone L (2010) Why do females have so few extrapair offspring? Behav Ecol Sociobiol. doi:10.1007/s00265-010-1104-z

Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it, and how did it evolve? Science 298:1569–1579

Henrich J, Gil-White FJ (2001) The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evol Hum Behav 22:165–196

Henrich J, Boyd R, Bowles S, Gintis H, Fehr E, Camerer C, McElreath R, Gurven M, Hill K, Barr A, Ensminger J, Tracer D, Marlowe FW, Patton JQ, Alvard M, Gil-White FJ, Henrich NS (2005) "Economic man" in cross-cultural perspective: ethnography and experiments from 15 small-scale societies. Behav Brain Sci 28:795–855

Holden CJ, Sear R, Mace R (2003) Matriliny as daughter-biased investment. Evol Hum Behav 24:99–112

Houston, A (2010) Central-place foraging by humans: transport and processing. Behav Ecol Sociobiol. doi:10.1007/s00265-010-1119-5

Huberman BA, Adamic AL (1999) Growth dynamics of the World-Wide Web. Nature 401:131

Kacelnik A, Vasconcelos M, Monteiro T, Aw J (2010) Darwin's 'Tug-of-War' vs. Starlings' 'Horse-Racing': how adaptations for sequential encounters drive simultaneous choice. Behav Ecol Sociobiol. doi:10.1007/s00265-010-1101-2

Kahneman D (2003) Maps of bounded rationality, psychology for behavioral economics. Am Econ Rev 93:1449–1475

Koerper HC, Stickel EG (1980) Cultural drift: a primary process of culture change. J Anthropol Res 36:463–469

Lachlan RF, Slater PJB (2003) Song learning by chaffinches: how accurately and from where? A simulation analysis of patterns of geographical variation. Anim Behav 65:957–969

Laherrère J, Sornette D (1998) Stretched exponential distributions in nature and economy, 'fat tails' with characteristic scales. Eur Phys J B 2:525–539

Laland KN (2004) Social learning strategies. Learn Behav 32:4–14

Larsen ON (1961) Innovators and early adopters of television. Sociol Inq 32:16–33

Lee R, DeVore I (1976) Kalahari hunter-gatherers. Harvard University, Cambridge, MA

Liljeros F, Edling CR, Amaral LAN, Stanley HE, Aberg Y (2001) The web of human sexual contacts. Nature 411:907–908

Lipo CP, Madsen ME, Dunnell RC, Hunt T (1997) Population structure, cultural transmission and frequency seriation. J Anthropol Archaeol 16:301–333

Lynch A, Baker AJ (1994) A population mimetics approach to cultural evolution in chaffinch song: differentiation among populations. Evolution 48:351–359

Marshall, J (2010) Ultimate causes and the evolution of altruism. Behav Ecol Sociobiol. doi:10.1007/s00265-010-1110-1

Marshall J, McNamara J, Houston A (2010). The state of Darwinian theory. Behav Ecol Sociobiol. doi:10.1007/s00265-010-1121-y

Mesoudi A, Lycett SJ (2009) Random copying, frequency-dependent copying and culture change. Evol Hum Behav 30:41–44

Neiman FD (1995) Stylistic variation in evolutionary perspective. Am Antiquity 60:7–36

Newman MEJ (2005) Power laws, Pareto distributions and Zipf's law. Contemp Phys 46:323–351

Ormerod P (2006) Why most things fail: evolution, extinction and economics. Pantheon Books, New York

Pareto V (1907) Manuel d'économie politique. Giard et Briére, Paris

Pagel M, Atkinson QD, Meade A (2007) Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. Nature 449:717–720

Perline R (2005) Strong, weak and false inverse power laws. Stat Sci 20:68–88

Powell A, Shennan SJ, Thomas MG (2009) Late Pleistocene demography and the appearance of modern human behavior. Science 324:1298–1301

Price DJD (1965) Networks of scientific papers. Science 149:510–512

Ratnieks FLW, Foster KR, Wenseleers T (2010) Darwin's special difficulty: the evolution of "neuter insects" and current theory. Behav Ecol Sociobiol. doi:10.1007/s00265-010-1124-8

Redner S (1998) How popular is your paper? An empirical study of the citation distribution. Eur Phys J B 4:131–134
Renfrew C, Scarre C (eds) (1998) Cognition and material culture. McDonald Institute for Archaeological Research, Cambridge, MA
Rogers EM (1962) Diffusion of innovations. Free Press, New York
Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. Science 311:854–856
Shennan SJ (2000) Population, culture history, and the dynamics of change. Curr Anthropol 41:811–835
Shennan SJ, Wilkinson JR (2001) Ceramic style change and neutral evolution: a case study from Neolithic Europe. Am Antiquity 66:577–594
Simon HA (1955a) A behavioral model of rational choice. Q J Econ 59:99–118
Simon HA (1955b) On a class of skew distribution functions. Biometrica 42:425–440
Slater PJB, Ince SA (1979) Cultural evolution in chaffinch song. Behaviour 71:146–166
Smith VL (2003) Constructivist and ecological rationality in economics. Am Econ Rev 93:465–508
Srinivasan V, Mason CH (1986) Nonlinear least squares estimation of new product diffusion models. Mark Sci 5:169–178
Steward J (1955) Theory of culture change. University of Illinois Press, Urbana, IL
Stigler GJ (1961) The economics of information. J Polit Econ 69:213–225
Stiglitz JE (2002) Information and the change in the paradigm in economics. Am Econ Rev 92:460–501
Weissing FJ, Edelaar P, van Doorn GS (2010) Adaptive speciation theory: A conceptual review. Behav Ecol Sociobiol. doi:10.1007/s00265-010-1125-7
Wiessner P, Schiefenhövel W (1997) Food and the status quest. Berghahn, Providence, RI
Winterhalder B, Smith EA (2000) Analyzing adaptive strategies: human behavioral ecology at twenty-five. Evol Anthropol 9:51–72
Yu DW, Sheperd GH (1998) Is beauty in the eye of the beholder? Nature 396:321
Yule GU (1925) A mathematical theory of evolution based on the conclusions of Dr. JC. Willis. Philos Trans R Soc B 213:21–87
Zipf GK (1949) Human behavior and the principle of least effort. Addison Wesley, Cambridge, MA