Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER Journal of Economic Dynamics & Control 31 (2007) 2438–2460
JOURNAL OF Economic Dynamics & Control
www.elsevier.com/locate/jedc

The emergence of Zipf's Law in a system of cities: An agent-based simulation approach

Yuri Mansury[a,b,*], László Gulyás[c,d]

[a] *Department of City and Regional Planning, Cornell University, Ithaca, NY 14853, USA*
[b] *KDI School of Public Policy and Management, Seoul 130-868, Republic of Korea*
[c] *AITIA International Inc., Budapest H-1039, Hungary*
[d] *Lorand Eotvos University, Budapest H-1117, Hungary*

Received 18 December 2004; accepted 15 August 2006
Available online 23 October 2006

Abstract

This paper develops a spatial agent-based model to generate a system of cities that exhibits the statistical properties of the Zipf's Law. The numerical results suggest that the combination of bounded rationality and maximum heterogeneity of agents can produce a generic power-law relationship in the size distribution of cities, but does not always generate the Zipf's Law. We found sufficient conditions on the probability distribution of spatial reach to generate the Zipf's Law associated with unit power coefficient. Our model also indicates that the Zipf's Law breaks down unless the extent of agglomeration economies overwhelms the negative disagglomerating forces.

JEL Classification: C61; C63; R12; P11

Keywords: Spatial agent based models; Power laws; Complex systems

*Corresponding author. KDI School of Public Policy and Management, 207-43 Cheongnyangni 2-Dong, Dongdaemun-Gu, Seoul 130-868, Republic of Korea. Tel.: +82 2 3299 1032; fax: +82 2 3299 1240.
E-mail address: ysm@kdischool.ac.kr (Y. Mansury).

0165-1889/$ - see front matter
doi:10.1016/j.jedc.2006.08.002

1. Introduction

The Zipf's Law embedded in the size distribution of US cities is one of the most-striking empirical regularities in urban economics. In brief, the Zipf's Law posits that if one ranks cities in descending order according to their population size, and then estimates the following equation:

$$\log(rank_i) = const + b \log(population_i) + \varepsilon_i \tag{1}$$

then one would get a straight downward-sloping line with the slope parameter b being very close to negative unity and a tight fit as measured by the coefficient of determination R^2. This type of linear association in the log–log scale belongs to the generic family of power-law relationships. If in addition the slope b equals negative unity, then such a special case of the power-law relationship is known to exhibit the Zipf's Law spatial pattern. Regardless of its magnitude, in this paper we designate the absolute value of the slope b as the power coefficient.

Auerbach (1913) is generally credited as the first to observe this empirical regularity in the size distribution of cities, which later on received greater recognition due to Zipf (1949), whom the "law" also owes its name to. Despite the early discovery, the quest for a robust theoretical model to explain such an empirical distribution of cities remains elusive; starting with Christaller (1933) who described Auerbach's finding as 'a most incredible law,' yet essentially was 'nothing more than just playing with numbers'. The view that the Zipf's Law is "atheoretical" carries on to the present date. Although numerous empirical studies have confirmed the existence of Zipf's Law in the size distribution of cities, the underlying mechanisms remain largely unexplored.

Previous empirical studies generally found that OLS estimates of Eq. (1) have statistical properties that are both stable and persistent. Stability refers to the magnitude of the power coefficient, $|b| \approx 1$, and the tight fit of such estimates, $R^2 \approx 1$, that have changed little over a long duration. For example, using 1997 US Census Bureau data, we obtain a power coefficient of 1.012 and $R^2 = 0.99$. A little further back, Krugman (1996) employed 1990 data and found a power coefficient of 1.003. It turns out that this stability property extends all the way back to the 1900 US Census (Dobkins and Ioannides, 2000), yielding remarkably similar results. Persistency, on the other hand, refers to the rankings of cities based on population size that have not changed very much. For example, the list of the ten largest metropolitan areas (MAs) was perfectly preserved between 1990 and 2000.[1] Indeed, using the US Census Bureau data we found that the rank correlation between the 1990 and 2000 size rankings for the entire MAs in the US is 0.99, indicating that there has been very little change during the intercensal years.

This study proposes a numerical model of interacting agents to generate a spatial pattern that is consistent with the observed statistical properties of the Zipf's Law. Such bottom-up approach connects the micro-behavior of individual agents to the

[1] The ten largest MAs in both years are: (1) New York, (2) Los Angeles, (3) Chicago, (4) Washington, DC, (5) San Francisco, (6) Philadelphia, (7) Boston, (8) Detroit, (9) Dallas, and (10) Houston.

emergence of the Zipf's Law at the aggregate level. Thus, the focus of this agent-based model is not on the formation of any single city, but on the spatial distribution of agents in a *system* of cities. To contrast our study with existing theoretical works on city systems that utilize analytical techniques to derive closed-form solutions (e.g., Simon, 1955; Simon and Bonini, 1958; Gabaix, 1999), we resort here to a numerical approach because local interactions in a heterogeneous population of agents render the model very difficult to solve analytically.

Our proposed model is ''agent based'' because the smallest unit of observation is an individual, which can be a firm or a household, rather than a group of aggregated individuals. To our knowledge, the use of agent-based models in economics began more than three decades ago with Cyert and March (1963) model of adaptive firms in a market system, and later was highlighted by Schelling (1978) seminal work on racially-segregated housing markets. With current advances in object-oriented computer programming, the use of agent-based model to study economic systems started to gain greater recognition, e.g., Alvin and Foley (1992) for modeling decentralized exchanges and LeBaron (2002, 2006) and Hommes (2006) for financial markets. The common aim in all these studies is to derive predictions of complex aggregate behavior from simple propositions governing the behavior of individual agents.

The idea that cities should be studied from the perspective of autonomous individuals has been proposed for a long time (see, e.g., Hagerstrand, 1970). However, only recently that there is a growing realization (Kirman, 1992) that building micro foundations to explain complex aggregate patterns has to start by explicitly considering local interactions among neighboring individuals. In this paper, such locally-interacting individuals represent ''firms,'' although they can be equally treated as consumers since the model does not distinguish between the demand and supply sides of the market. We adopt such simplification because it allows our study to focus on the main drivers of a migrating population that can generate the Zipf's Law within a spatial model without the complications of a fully-specified market system.

Our model does not rely on location-specific environmental characteristics to explain migration pattern. Instead, agents evaluate the attractiveness of spatially homogeneous locations based on pure demographics. Specifically, agents gravitate towards high-density locations because such locations generate positive externalities, which have been shown empirically to be the driving force behind the formation of agglomeration economies in the US (see, e.g., Rosenthal & Strange, 2001). The formation of a city system in our model therefore results from an endogenous process that emerges from the internal logic of a self-organizing system. Hence we abstract away from external sources of spatial agglomeration.

Our main findings can be summarized as follows. Our agent-based model generates robust Zipf's Law under the following set of conditions. First and foremost, we found that to produce a power-law relationship in the size distribution of cities requires bounded rationality in the form of limited spatial reach. Bounded rationality alone, however, is not sufficient to replicate the statistical properties of the Zipf's Law (i.e. with both the power coefficient b and R^2 from OLS estimates of

Eq. (1) approaching unity). With the help of numerical simulations, we identified sufficient conditions on agents' spatial reach that can generate a system of cities whose rank-size distribution follows the Zipf's Law. In Section 2, we review previous theoretical models of the endogenous formation of city systems.

2. Previous studies

2.1. Models of city formation

One strand of theory proposes that cities were formed as "hubs" to minimize the firms' costs of transporting raw materials to their factories and of transporting their products to the consumers (see, e.g. Isard, 1952; Alonso, 1964). The contemporary US economy, however, is increasingly dominated by the service industry that has minimal tangible products to transport. As Weber (1914) pointed out, firms often agglomerate for reasons other than to minimize transport costs. Instead, the agglomerating force can simply be the sheer size of a city as reflected by its population density. A number of theoretical arguments have been advanced in the attempt to explain "why size matters." For example, according to the argument developed by Mills (1967) and Henderson (1974), cities emerge because of economies of scale. Specifically, firms like to have other firms nearby because of the complementary nature of their production processes, or because being together gives them the ability to attract customers to a one-stop shopping location. Henderson (1986) subsequently found the empirical evidence that the extent of external economies of scale is indeed significant among US cities. The Henderson–-Mills framework nonetheless lacks explicit spatial consideration.

In recent years, urban economists have increasingly relied on numerical methods to investigate the spatial dynamics of a system of cities.[2] For example, using a 1D spatial model, Krugman (1996) focuses on the tension between attraction and repulsion (centripetal and centrifugal forces in his terminology) that ultimately creates multiple cities. As in here, Krugman postulates that the endogenous formation of cities is driven by the self-organizing behavior of migrating agents. His model, however, does not allow explicit decision-making at the individuals' level.[3] More recently, Page (1999) proposes a computational model to understand the formation of cities based on individual migration decisions that are driven by population dynamics. None of such spatial computational models, however, seeks to explain the emergence of Zipf's Law based on the agglomerating effects of positive externalities.

2.2. Models that generate Zipf's rank-size distribution of cities

A growing body of literature is devoted to explain the emergence of power laws. Newman (2005) summarizes the best known and most widely applied processes that

[2]See the review in Fujita et al. (1999).

[3]Gulyás (2002) proposes a modification of the Krugman's model that does allow for individual-level decision-making.

have been proposed to generate power laws. It turns out that many of these models generate a power law distribution by assuming another one, yet does not explain the origin of the underlying power-law relationship. Thus, although by no means trivial, the results are hardly surprising. Other models require pure random walk processes, which hardly correspond to human's migratory behavior that is biased towards certain locations. An exception is the mechanism of preferential attachment known as the Yule process, which later inspired the works of Herbert Simon. The Yule model can generate a wide range of power coefficients (including one, i.e. the Zipf's Law), but is not suitable to model city formation because agents can never relocate once they are attached to a location. Beyond the power laws, the Gibrat (1931) principle of proportionate effect generates a log-normal distribution, which approximates the power law in a small segment of the curve. Unlike the power law distribution, however, a log-normal distribution has a finite mean and variance (see Mitzenmacher, 2004), and its full plot in a log–log chart represents a quadratic relationship rather than a straight line.

Theoretical studies that examine the distribution of population among a group of cities can be classified as those utilizing hierarchical or random-growth models. The origin of the former can be traced to the central place theory of Christaller (1933), and its early formal exposition can be found for example in Beckman (1958) and Parr (1970). Hierarchical models are useful as a data descriptive tool to characterize cities of various sizes as a connected spatial network with central places at the top of the hierarchy. However, such models do not explain the mechanism by which individual migration decisions can lead to the Zipf's Law spatial pattern.

In contrast, random-growth models rely on the mechanism of preferential attachment to generate the Zipf's Law (see Simon (1955) and Simon and Bonini (1958). See also Gabaix (1999) for a recent contribution). The common feature of all these models is the appearance of the Zipf's Law in the stationary state of a stochastic process. In particular, Simon proposed a stochastic process that begins with the arrival of new agents in every period. An agent belonging to this batch then promptly attaches itself to an existing location with a probability that is proportional to the population size of that location. Although Simon's approach provides important insights into the underlying stochastic process that generate the Zipf's Law, it requires perpetually growing number of agents that permanently attach themselves to their initial locations. Further, these agents are not autonomous individuals since they follow a probabilistic rule of attachment that is imposed in a "top-down" manner.

A number of theoretical models of city formation have also appeared in physics journals. For example, Zanette and Manrubia (1997), Manrubia and Zanette (1998), and Makse et al. (1998) proposed models that combined stochastic events and deterministic diffusion processes to explain the power-law relationship in a system of cities. Their models are spatial, but not based on bottom-up, individual level behavior. Specifically, in the model of Zanette and Manrubia (1997, 1998), the elementary unit of observation is a city whose population dynamic is assumed to follow a stochastic path characterized by intermittency. Makse et al. (1998) applied insights from percolation theory to reproduce urban-growth dynamics characterized by power-law relationship. This model assumes perfect information structure, and agglomeration emerges from agents attaching themselves to a cluster according to a top-down stochastic rule as in the preferential attachment model of Simon (1955) and Simon and Bonini (1958). All these studies focus on the probability masses instead of on the individuals that make up these masses.

An alternative way is to view a system of cities as a network of industrial activities and social ties. Watts (1999) has argued that many social networks behave as "small worlds" characterized by simultaneous local clustering and global connectivity. A common property of such networks is that the connecting links follow a power law distribution (see Strogatz, 2005). To generate such distribution, however, existing network models assume a preferential attachment mechanism similar to that of Simon (see Barabasi and Albert, 1999; Menczer, 2004). Further, a network model typically neglects the role of space and therefore is not suitable for modeling the migratory behavior of agents in a city system.[4] We note that exceptions do exist, including the work of Andersson et al. (2003) that introduces a spatial dimension to the original Barabasi–Albert network model to generate a power-law distribution of urban land prices. Their model, however, is not designed to explain the rank-size distribution of *human population* (i.e. the Zipf's Law). Nonetheless, the application of small world theory in studies of an urban system remains a fertile area for future research. As Batty (2001) points out, a properly-developed network theory for cities can very well hold the key to a better understanding of the complex connectivity in and between cities.

In the following section, we present our agent-based model in which it is the migratory behavior of autonomous decision-makers that drives the formation of large-scale spatial patterns at the macro level.

3. Mathematical modeling

The spatial pattern of population distribution in our model changes due to agents' decision to migrate. To fix ideas, let $S_i(t)$ denote the state of location i at time $t = 1,2,\ldots, T$ whose value is updated according to the following rule:

$$S_i(t+1) = f(S_i(t), S_{\Omega(i,\delta)}(t)), \tag{2}$$

where f is a real-valued function $f : \Re \times \Re^{\mathbf{\Omega}} \to \Re$, and Ω (i,δ) is the set of neighbors that are within δ units of distance away from location i: Ω (i,δ) $= \{j \neq i : d_{ij} \leqslant \delta\}$. Eq. (2) specifies that locations change their states at rates that depend not only on their own previous states, but also on the previous states of the neighboring sites within a fixed distance. To describe the dynamics of migration, let $A(t)$ denote the set of all agents in the system at time t, and suppose every agent $a \in A(t)$ is characterized by its capability to migrate to locations that are within a certain distant threshold, $r_a \geqslant 1$. Then the migratory decision of an agent depends on the state of its current site as

[4] Watts and Strogatz (1998) adds a spatial dimension to their network model, but the authors did not show whether their spatially-extended model retains the capability to generate a power-law relationship.

well as that of the neighboring locations that it can reach through migration

$$P_a(t+1) = g\left(P_a(t), S_{P_a(t)}, S_{\Omega(P_a(t),r_a)}, \mathbf{E}(t)\right), \tag{3}$$

where P_a (t) represents the spatial position of agent a, Ω $(P_a(t),r_a)$ the set of neighboring locations that the agent can reach, and $\mathbf{E}(t)$ a vector of environmental variables. For example, $S_i(t)$ could represent the attractiveness of location i in quantitative terms, while $\mathbf{E}(t)$ a vector of environmental factors that are exogenous to the agents (e.g. energy, infrastructure, natural resources, pollution). In this study, we assume a spatially homogeneous environment with the aim to show that population dynamics alone are sufficient to drive changes in P_a (t) and $S_i(t)$ which, in turn, generate the Zipf's Law. The detailed specifications of our agent-based model are as follows.

3.1. Spatial aspects of the model

Our space of observation is a grid lattice that comprises a countable collection of Z^2 sites, where Z is the finite length of the two-dimensional grid. To assign unique identification, each location is indexed by the subscript $i = x_iZ+y_i$, with $(x_i, y_i) \in [0\ldots Z–1] \times [0\ldots Z–1]$. To quantify the distance between any two locations i and j, d_{ij}, we employ the l-infinity metric that defines $d_{ij} =$ Max [abs $(x_j–x_i)$, abs $(y_j–y_i)$].

3.2. Population dynamics

At any given time, location i can be either empty or populated by one or more agents. Let $n_i(t) \geqslant 0$ represent the population size of location i at time t, and $N(t) \equiv \sum_i n_i(t)$ the total number of agents in the entire system. At the beginning of time, we introduce $\bar{N}$ number of agents into the world. As in Gibrat (1931), we assume infinitely lived agents that bear no offspring, implying a fixed total population for all periods, $N(t) = \bar{N}, \forall t = 1, 2, \ldots, T$. Each agent $a \in A(1)$ is randomly assigned an initial location based on a uniform probability distribution function (PDF).

A city is defined as any location i hosting at least a single agent, $n_i(t) > 0$. With a fixed population size in the entire system, the population dynamics of cities in our model are fully determined by net migration flows. Formally, denoting $\gamma_{ij}(t)$ as the number of agents migrating from location i to j, the size of a city formed in location i at time t can be computed as

$$n_i(t) = n_i(t-1) + \sum_j \gamma_{ji}(t-1) - \sum_j \gamma_{ij}(t-1). \tag{4}$$

That is, the population size of a city changes if the previous flows of in-migration, $\sum_j \gamma_{ji}(t-1)$, are not counterbalanced by the flows of out-migration, $\sum_j \gamma_{ij}(t-1)$.

3.3. Search process

Agents exhibit bounded rationality in the sense of Simon (1956), which is here manifested in agents' inability to assess locations outside of their maximum visibility reach, v_a. Search is therefore *local* because an agent can only identify more attractive locations that are within its visibility range.[5] This incomplete information structure suggests that agents are optimizing under spatial constraints, which can be viewed as technological upper bounds prohibiting agents' access to information contained in out-of-reach locations, or the increasing costs associated with relocation to faraway places.[6]

At the beginning of time, every agent a is randomly assigned a maximum visibility reach $v_a \in [v_{\min}, v_{\max}]$, where $1 \leqslant v_{\min} \leqslant v_{\max} \leqslant Z/2$. The bandwidth of the PDF domain, $v_{\max} - v_{\min}$, measures the extent of agents' heterogeneity in our model. To see this, note that wider bandwidth implies greater variability of the spatial reach distribution as long as $p(v_a) > 0$ for $v_a \in [v_{\min}, v_{\max}]$. Limited visibility reach in turn translates into limited mobility preventing agents to migrate to locations that are not visible to them

$$r_a \leqslant v_a, \tag{5}$$

where r_a denotes agent a's spatial reach corresponding to the maximum distance that the agent can traverse within a single period. Agents with greater vision therefore are more mobile because in a single step they can migrate to locations that their peers with more limited vision cannot. Given its maximum visibility reach, we assume that agent a can only migrate to locations that are within $r_a = v_a/2$ units of distance away, thus ensuring that the neighborhood size is the same for every location within the spatial reach of the evaluating agent. For example, $r_a \in [1, Z/2]$ if within one period the least-mobile agents can only migrate to their immediate neighbors while the most-mobile agents can reach every location in the entire system.

3.4. Migratory behavior

The equation governing migratory behavior remains to be specified. It is well known that the location decisions of individual firms are heavily influenced by the presence of *agglomeration economies* (see, e.g. Ellison and Glaeser, 1997). Given the set of locations that agent a can reach, $\Omega(P_a(t), r_a)$, we assume that the attractiveness of location $j \in \Omega(P_a(t), r_a)$ depends on the extent of on-site agglomeration economies. Specifically, here we adopt the following additively-separable linear specification:

$$\Psi_{j,a}(t) = n_j(t) - c \cdot n_j(t)^2, \tag{6}$$

where $\Psi_{j,a}$ represents the value of location j to agent a, while c captures the negative impact of population overcrowding. Note that $\Psi_{j,a}$ is agent specific because it depends not only on the evaluating agent's current position but also on her spatial reach. The first two terms on the right-hand side of Eq. (6) represent agglomeration economies as a function of location j's population density, n_j. This agglomeration

[5]Note, however, that Simon in his 1955 paper implicitly assumes global search because there every agent is aware of all the existing locations in the system.

[6]Alternatively, the source of bounded rationality could be due to costly "deliberation" (see the survey article of Conlisk, 1996) or, as MacLeod (2002) suggests, due to the rising "complexity" of the problem beyond a certain *threshold*, which in here is represented by agents' maximum visibility reach.

effect can be broken down further into positive externalities (the first term) and the negative effect of population overcrowding (the quadratic second term).

The source of positive externalities as captured by the first term has been the subject of numerous theoretical endeavors.[7] For example, Murphy et al. (1989) and Krugman (1991) have argued that geographic concentration of firms brings about physical spillovers vis-à-vis lower costs of infrastructure. On the other hand, Glaeser et al. (1992) suggest that agglomeration economies generate intangible spillovers of knowledge and ideas to the neighboring firms, which raise the average productivity of all firms within the geographic proximity and hence reinforce agglomeration further. Empirically, Roback (1982) shows that population density itself can directly represent a desirable amenity in the sense that it has a positive imputed price. Henderson (1986) presents the empirical evidence showing that firm productivity is higher in locations where there are neighboring firms from the same industries.

In contrast, the negative second term in Eq. (6) represents the disagglomerating effect of highly populated regions. Specifically, a high-density industrial center is often associated with higher levels of crime rates, pollution, land costs and general costs of living. This kind of external *dis*economies can lead to out-migration of residents that seek to avoid highly congested areas. For example, Cullen and Levitt (1999) have shown that on average, a 10 percent increase in crime rates subsequently leads to a one percent decline in population. Here in Eq. (6) the parameter c captures the extent of these disagglomerating forces. Smaller values of c corresponds to a higher threshold level of population, $n_j = 1/(2c)$, which marks the point where the negative disagglomerating effect starts to overwhelm the positive impact of agglomeration economies as location j expands further, $\partial\Psi_{j,a}/\partial n_j\big|_{n_j>1/(2c)} < 0$.

Once the value of every visible location has been calculated, in the next period agent a migrates to the location that from his point of view yields the highest value[8]

$$P_a(t+1) \equiv \arg\ \max_{j\in\Omega[P_a(t),r_a]} \Psi_{j,a}(t). \tag{7}$$

Eq. (7) shows that despite the spatial constraints imposed by their limited visibility and spatial reach, nevertheless agents could still select the best location available *within* their limited information set. Thus, in contrast to models of bounded rationality that tend to dispense entirely with the notion of optimization (e.g. Nelson and Winter, 1982), here our notion of limited spatial reach does not preclude optimizing procedures, but without the need for every agent to exhibit perfect foresight.

Our discrete model has so far eluded our attempts to derive analytical solutions. Therefore, we resorted to numerical simulations to examine the statistical properties of the cities' rank-size distribution. The following section details the algorithm of the agent-based simulation model.

[7]We are aware of the difference between positive externalities that promote city growth (e.g. Glaeser et al., 1992), and those that promote city formation (e.g. Henderson, 1986). In our model, however, these two effects yield the same result.

[8]In the case of multiple maxima, the actual location chosen is randomly determined with equal probabilities.

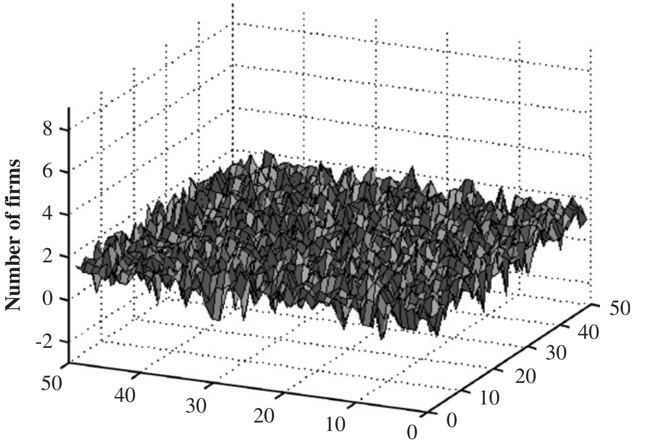

Fig. 1. Initial, random-uniform distribution of agents. The X–Y axes correspond to the spatial coordinates of various locations on the grid lattice, while the Z-axis represents the population size of these locations.

4. Algorithm implementation

At the beginning of every simulation reported here, we place 12 000 agents uniformly randomly on a grid lattice with length Z as shown in Fig. 1. The edges of this space lattice are joined in such a way that locations on opposite edges are connected. The formation of small, initial settlements can be thought of as due to geographic characteristics or resource endowments that render certain locations more attractive than the other (e.g. coastal proximity as in Rappaport and Sachs, 2003). We assume that no agents are completely immobile: $r_a \geqslant 1$ for all $a \in A$, so that the least mobile agent can reach locations that are one unit of distance away.

The spatial reach is distributed randomly among agents based on a discrete probability distribution function (PDF) with finite support, $r_{\min} \leqslant r_a \leqslant r_{\max}$. Thus, within a single period the least-mobile agents can only reach locations that are $r_{\min}$ units of distance away, while the most-mobile agents can possibly travel from one end of the world to the other, i.e. $r_{\max} \leqslant Z/2$. The various discrete PDFs of r_a that we have experimented with here all can be easily approximated by the Beta distribution, which comprises all continuous PDFs of the form

$$p(\hat{r}_a) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}(1-\hat{r}_a)^{\beta-1}\hat{r}_a^{\alpha-1}, \tag{8}$$

where $\hat{r}_a = [r_a - E(r_a)]/Var(r_a)$ standardizes the spatial reach such that $0 \leqslant \hat{r}_a \leqslant 1$, $\Gamma(\cdot)$ the Gamma function, and $\alpha > 0$, $\beta > 0$ constant parameters.[9] By varying these two parameters, α and β, the Beta distribution allows us to approximate the various discrete PDFs of r_a actually used in the simulations using a known continuous form. For example, by setting $\alpha = \beta$ we can obtain the family of symmetric distributions,

[9]For a complete treatment of the Beta PDF, see e.g. Johnson et al. (1995).

which includes the uniform and normal distributions as special cases. Alternatively, decreasing α so that $\alpha < \beta$ skews the PDF to the right, representing the case where there is greater likelihood to find agents with less-than-average mobility. In contrast, increasing α so that $\alpha > \beta$ skews the PDF to the left corresponding to a system in which there is greater likelihood to find agents with higher-than-average mobility.

For a given grid length Z and the support $r_{\min} \leqslant r_a \leqslant r_{\max}$, we can describe the discrete PDF of agents' spatial reach using a series of positive integers, $\{w_r\} = w_{r_{\min}}, w_{r_{\min+1}}, \ldots, w_{r_{\max}}$, each representing the probability weight of the corresponding $r_a \in [r_{\min}, r_{\max}]$. For example, given the grid length $Z = 10$ and the support $1 \leqslant r_a \leqslant 5$, then a uniform PDF can be fully described by the series $\{w_r\}$ = "1 1 1 1 1", which means $\mathrm{p}(r_a \in [1, \ldots, 5]) = w_{r_a} / \sum_{r=r_{\min}}^{r_{\max}} w_r = \frac{1}{5}$. For convenience, we use the "*" character to represent a consecutive series of unit weights. Thus, the same uniform PDF of agents' spatial reach $\mathrm{p}(r_a)$~"1 1 1 1 1" can be equivalently described as $\mathrm{p}(r_a)$~"*".

Non-uniform PDFs necessarily involve various weights, $w_r \geqslant 1$. If the weights are strictly greater than one, $w_r > 1$ for all $r_{\min} \leqslant r_a \leqslant r_{\max}$, then such series can be normalized by assigning unit weight to the spatial reach with the smallest positive probability. Thus, consider the asymmetric PDF with the support $1 \leqslant r_a \leqslant 5$ for example, $\mathrm{p}(r_a)$~"8 2 2 2 2". This right-skewed distribution can be equivalently described by the normalized PDF $\mathrm{p}(r_a)$~"4 *"≡"4 1 1 1 1", since both imply $\mathrm{p}(r_a = 1) = \frac{1}{2}$ and $\mathrm{p}(r_a \in [2, \ldots, 5]) = \frac{1}{8}$. We note that more than one asterisk may be included in a description of a non-uniform PDF, in which case the unit weights are simply distributed evenly among the asterisks. For example, given $Z = 14$ and the support $1 \leqslant r_a \leqslant 7$, the distribution $\mathrm{p}(r_a)$~"* 5 *" is equivalent to "1 1 1 5 1 1 1", implying $\mathrm{p}(r_a = 4) = \frac{5}{11}$ and $\mathrm{p}(r_a \neq 4 \cap r_a \in [1, \ldots, 7]) = \frac{1}{11}$.

At the end of a simulation, we compute the population size of all non-empty locations in the grid, rank them in descending order, and then estimate Eq. (1). The PDFs that are capable of replicating the empirical Zipf's Law pattern of US cities are designated as "admissible candidates."

Definition. A PDF of agents' spatial reach is an *admissible candidate* if the resulting power coefficient $b \in [0.95, 1.05]$ and the coefficient of determination R^2 from OLS estimates of Eq. (1) is greater than 0.95.

In our simulations, the starting point is to examine the barebone case where we set the parameter $c = 0$, i.e. no disagglomerating forces, and which is then followed by the case where c is non-zero. For each PDF of agents' spatial reach, we then vary the grid length Z, and for each set of parameter values we perform 30 Monte Carlo simulations with different random seeds to establish the robustness of our results.

5. Results

Our simulation results can be briefly summarized as follows. In all the simulations reported here, stable and persistent size distribution of cities emerges typically in less than 20 time periods, indicating rapid convergence into equilibrium.[10] With no restrictions imposed on the power coefficient, robust power-law relationships are yielded if the system contains sufficiently heterogeneous population of agents. However, to go from the generic power-law relationship into the specific case of the Zipf's Law requires further restrictions on the underlying PDF. In particular, we found that right-skewed and bimodal PDFs can successfully generate unit power coefficient (i.e. the Zipf's Law). Finally, we found that the Zipf's Law breaks down unless the extent of positive agglomeration economies significantly outweighs that of the negative disagglomerating forces.

5.1. The barebone case: $c = 0$, no disagglomeration forces

5.1.1. Initial explorations: $Z = 50$

To begin our exploration, we fix the grid length at $Z = 50$. With identical agents, $r_a = \bar{r}$ for all $a \in A$, it is straightforward to show (see, e.g. Krugman, 1996) that the result converges to a system of multiple, equidistant cities of similar size. In the limit, when every agent in the system has perfect vision, $r_a = \bar{r} = Z/2$, the entire population congregates in a single giant city. Thus, a model featuring identical agents in a spatially homogeneous environment has no chance of generating a system of multiple cities with unequal sizes. With that in mind, next we examine the rank-size distribution of cities that results from widening the bandwidth of the spatial reach domain $r_{\max}-r_{\min}$, while maintaining the parameter values $\alpha = \beta = 1$ of the Beta PDF (see Eq. (8)). But first, here we introduce the notion of "completeness."

Definition. The PDF of agents' spatial reach is *complete* in a system with grid length Z if the probability of finding an agent with spatial reach r_a is strictly positive, $\mathrm{p}(r_a) > 0$, for $r_a \in [1, Z/2]$, and zero otherwise.

Note that a complete PDF implies that the spatial-reach bandwidth equals $Z/2$–1, thus yielding maximum heterogeneity of agents defined in terms of their spatial reach. We found that such complete PDF guarantees a system of cities exhibiting a robust power-law relationship. This we can show by varying the lowest and highest reach of agents, $r_{\min}$ and $r_{\max}$, respectively. Fig. 2 shows the results of varying $r_{\max}$ while fixing the minimum spatial reach at its lower bound, $r_{\min} = 1$. When $r_{\max}$ exceeds ten, the power-law distribution of cities begins to emerge, as indicated by the straight downward-sloping line in the log–log scale that spans at least three orders of magnitude. However, lowering $r_{\max}$ while fixing $r_{\min} = 1$ breaks down the power-law relationship starting at $r_{\max} = 10$, resulting instead in a concave, downward sloping curve. To see why the power-law relationship breaks down, consider the formation

[10]Due to this rapid convergence, unlike in the Gibrat's model as Gabaix (1999) detailed, our model does not exhibit a period where cities all grow at a scale-invariant (non-zero) growth rate with a common mean. Obviously, once the system reaches equilibrium then the population growth rates and variance become trivially zero across all cities. This is due to the 'closed' system nature of our model with a fixed total population.

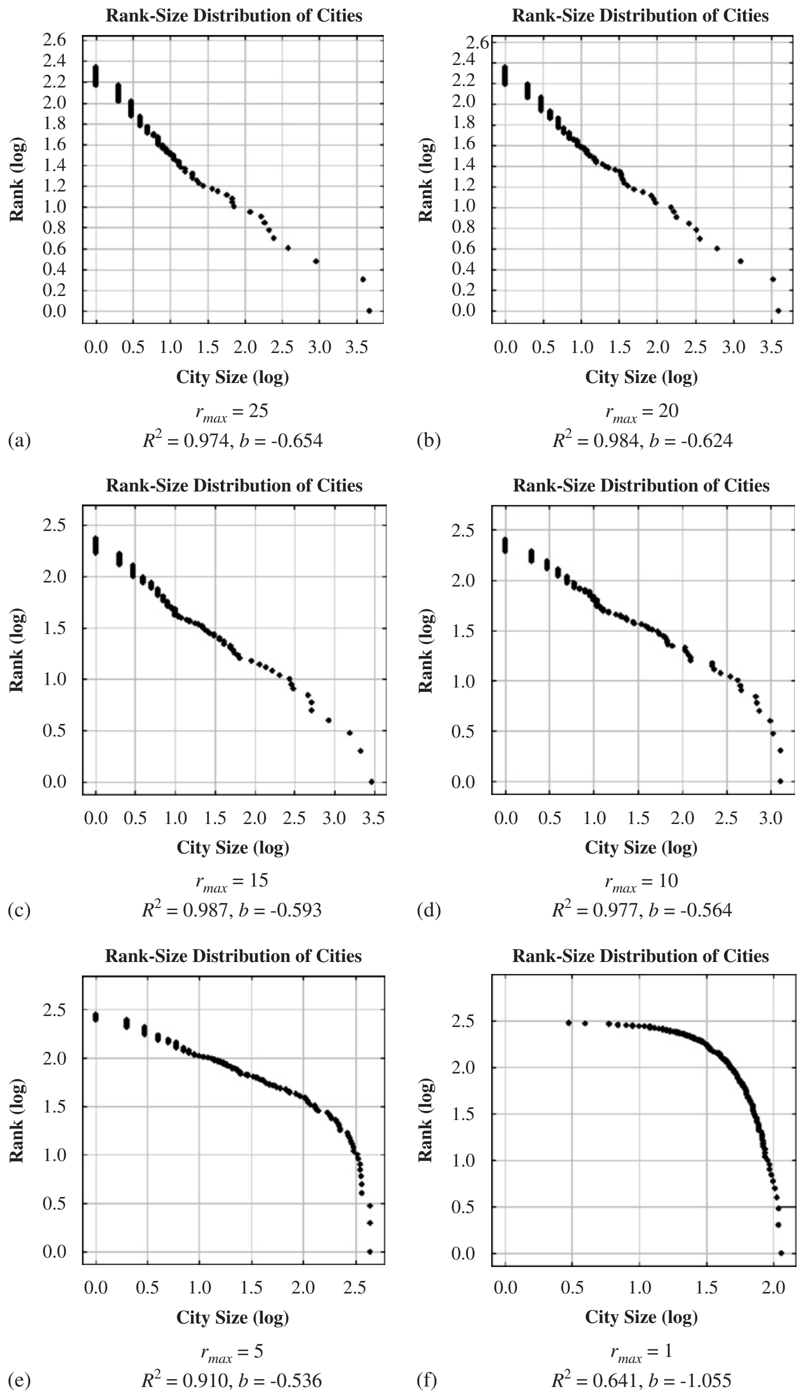

Fig. 2. Rank-size distribution of cities, various $r_{\max}$, fixed $r_{\min} = 1$. The X-axis corresponds to the population size of cities, while the Y-axis to their rank, in log–log scale. Shown also are the power coefficient b and the coefficient of determination R^2 from OLS estimates of Eq. (1).

of the largest city C. As initially agents are uniformly distributed across the landscape, the majority of C's residents must be agents migrating from other regions after successive multi-period journeys. For a typical migrating agent, the population densities of the different cities visited before it arrives at C form a monotonically increasing series. Lower $r_{\max}$ implies greater likelihood for an agent to terminate its journey early, which means fewer agents will ever make it to C, in turn eliminates extremely large cities rendering a more equal size distribution of cities, hence the breakdown of the power-law relationship.

If instead the maximum reach is fixed at the upper bound, $r_{\max} = Z/2$, then the power law relationship breaks down if we raise agents' minimum reach closer to the maximum, $r_{\min} \to r_{\max}$, because of the disappearance of smaller cities. A complete PDF with $r_{\min} = 1$ and $r_{\max} = 25$ for $Z = 50$ therefore guarantees the emergence of a power-law relationship in the size distribution of cities. Fig. 3 shows the resulting equilibrium system of cities when the underlying PDF is complete representing maximum heterogeneity of agents. The power-law relationship is due to a few mega-cities coexisting with many smaller settlements.

The initial explorations thus reveal that a complete PDF is sufficient to generate a power-law relationship in the size distribution of cities. However, although the OLS fit is satisfactory ($R^2 = 0.97$), the estimated power coefficient ($b = 0.65$ in absolute value) is significantly lower than one associated with the Zipf's Law. Since such empirical Zipf's Law relation with unit power coefficient is an instance of the generic power law, from here on our strategy is to focus only on complete PDFs. For comparison purposes, we designate p(r_a)~"*" as our *benchmark* distribution. That is, in our benchmark simulation r_a follows a uniform and complete PDF such that $p(r_a) = 1/(Z/2)$ for $1 \leqslant r_a \leqslant Z/2$, and zero otherwise.

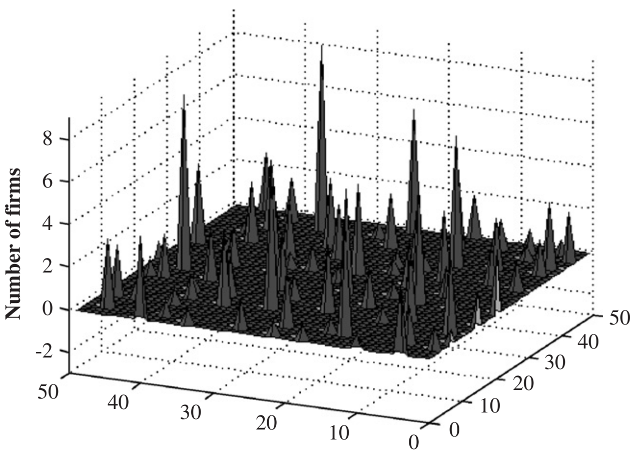

Fig. 3. Equilibrium spatial distribution of agents across cities that exhibit a power-law relationship. As shown, $Z = 50$, $r_{\max} = 25$, $r_{\min} = 1$, and p(r_a)~"*".The X– Y axes correspond to the spatial coordinates of various locations on the grid lattice, while the Z-axis represents the population size of these locations.

5.1.2. Robust simulation results

An exhaustive parameter search is impossible due to the astronomical number of complete PDFs for a given grid length Z. To see this, consider the general form of a complete PDF for $Z = 50$, whose domain for spatial reaches is a 25-tuple integer values (1, …, 25). Given $\bar{N} = 12\,000$, then the number of possible reach-configurations is $(Z/2)^{\bar{N}} = 25^{12000} \approx 10^{16696}$. Therefore, we limit our numerical explorations to a few well-known PDFs, namely the right skewed, left skewed, normal (Gaussian), "1 X * *", and bimodal distributions. The particular numerical values associated with these distributions are detailed below.

Fig. 4 compares the uniform PDF of r_a with non-symmetric PDFs that are right-skewed, namely "2 *", "5 *", "10 *", and "15 *, which assigns higher probability mass to the least-mobile agents. For example, in distribution "2 *" it is twice more likely to find myopic agents with $r_a = 1$ than their more mobile peers. This class of discrete distributions can be approximated by a continuous Beta PDF with $\alpha \leqslant 1 \ll \beta$. As Fig. 4 shows, our simulated results are robust across different random seeds as indicated by the small error bars for the benchmark PDF "*". However, if the purpose is to replicate the Zipf's Law spatial pattern, then the right-skewed PDFs cannot improve the results generated by the uniform distribution "*". Raising the weights attributed to the least-mobile agents results in power coefficient b being closer to unity (Fig. 4). Doing so, however, reduces the fit R^2. If we apply a strict set of admissibility criteria such that both the power coefficient b and the fit R^2 must be greater than 0.95, then the best outcome appears to be generated with the PDF "15 *" at $Z = 100$.

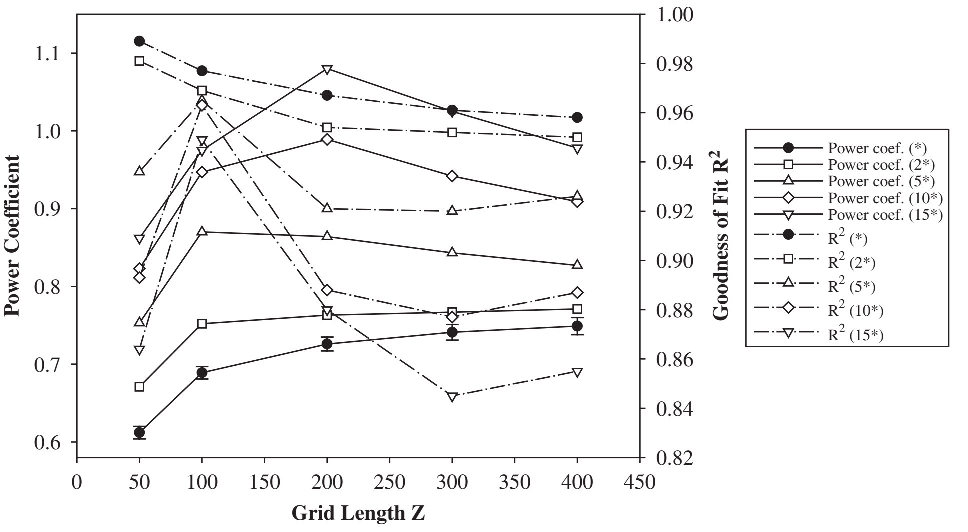

Fig. 4. Power coefficients b for the uniform and right-skewed 'X *' reach PDFs, various grid length Z. The left Y-axis represents the power coefficient (in absolute values); the right Y-axis the coefficient of determination R^2. For clarity, the error bars from the 30 Monte Carlo simulations using different random seeds are only shown for the benchmark, uniform distribution.

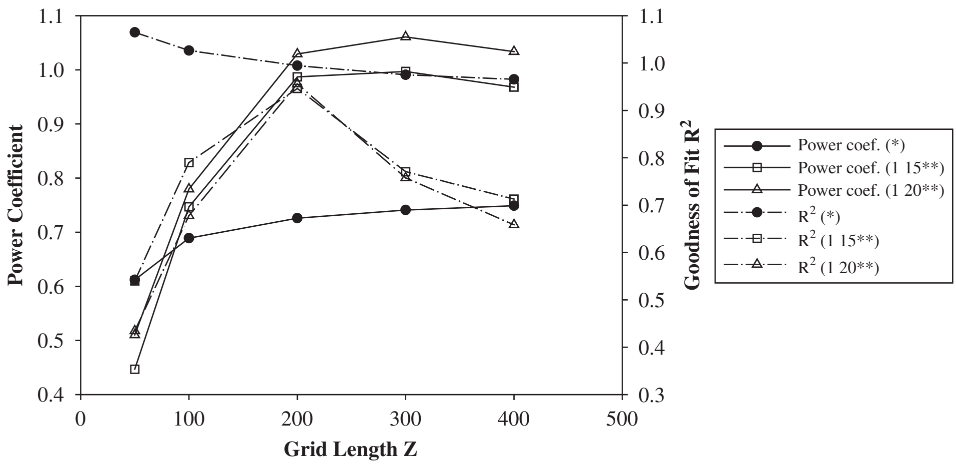

Fig. 5. Power coefficients b for the uniform and right-skewed '1 X * *' reach PDFs, various grid length Z. The left Y-axis represents the power coefficient (in absolute values); the right Y-axis the coefficient of determination R^2.

Next, we experimented with PDFs that are left-skewed, namely "* 2", "*5", "*10", and "* 15", which relate to the Beta distribution with $\alpha \geqslant 1 \geqslant \beta$. These left-skewed PDFs assign greater likelihood for an agent to be of the most-mobile type since the modal spatial reach is larger than the median, which, in turn, is larger than the mean. We found that, for a given grid length Z, left-skewed PDFs result in the power coefficient b and the fit R^2 that are similar to those of the benchmark uniform PDF, and in some cases with a slight tendency for the power coefficient of the former to be lower. Other well-known PDFs we experimented with include symmetric, single-peaked distributions, e.g. normal Gaussian, and those PDFs whose peaks are either at the second quartile, e.g. the "* 15 * *" distribution, or at the third quartile, e.g. the "* * 15 *" distribution. We obtained results that are not only inadmissible, but also inferior to the benchmark "*" distribution due to their lower values of the power coefficient b and lesser fit R^2.

It turns out that assigning greater weights to the second-lowest spatial reach, i.e. as in the "1 X * *" distribution where the integer $X > 1$ corresponds to the mode, results in significant improvement in the power coefficient b relative to that of the benchmark PDF "*". We can approximate these right-skewed PDFs using the continuous Beta distribution with $1 < \alpha \ll \beta$. Applying our selection criteria, we found distributions "1 15 * *" and "1 20 * *" (Fig. 5) to be admissible candidates. For example, at $Z = 200$, with these PDFs the power coefficient slope b can be raised to unity without compromising the fit R^2.

We also found robust Zipf's Law relationship if the PDFs exhibit bimodal peaks at both the lower and upper tails represented by the "X * X" distributions. Specifically, we found the bimodal PDFs represented by "10 * 5", "10 * 10", "10 * 15", "10 * 20", "15 * 5", "15 * 10", "15 * 15", and "15 * 20" to

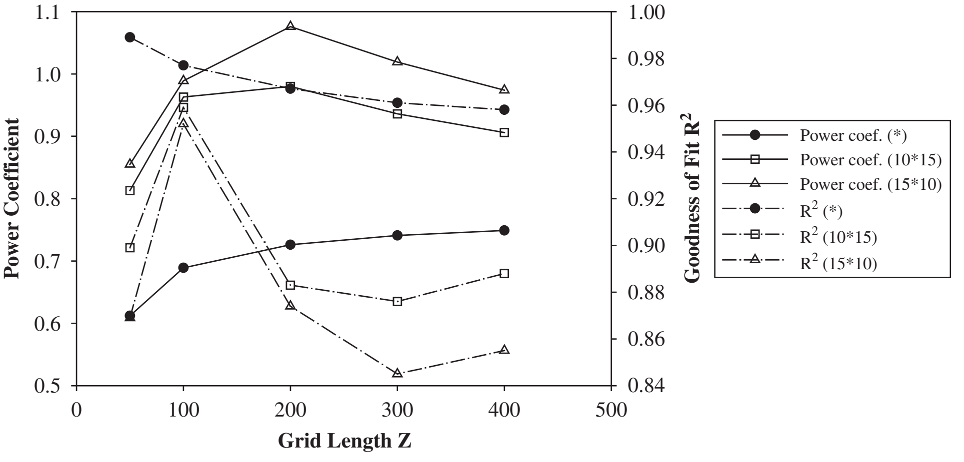

Fig. 6. Power coefficients b for the uniform and bimodal "$X * X$" reach PDFs, various grid length Z. The left Y-axis represents the power coefficient (in absolute values); the right Y-axis the coefficient of determination R^2.

be admissible candidates satisfying our selection criteria. Such bimodal PDFs can be approximated by a Beta distribution with $\alpha = \beta < 1$. Fig. 6 shows the results from selected bimodal PDFs (i.e. "10 * 15" and "15 * 10"), which for example can replicate the empirical Zipf's Law pattern of US cities at grid length $Z = 100$.

5.2. The inclusion of disagglomeration forces ($c > 0$)

In this section we explore the marginal impact of the disagglomeration parameter, c. In general, we found that rising disagglomeration effect results in lower power coefficients, thus moving away from the admissibility criterion because city sizes become more uniform with higher values of c. Specifically, Fig. 7 shows that, at $Z = 50$ and with the benchmark distribution "*", the power-law relationship breaks down when c is raised to a value higher than 0.035 as evident by the rank-size plot that is no longer a straight downward-sloping curve spanning three orders of magnitude in the log–log scale. The breakdown is due to a more equal spatial distribution of agents, as can be shown by the declining variance of the size of the six largest cities when the value of c is increased to a level higher than 0.035. Our finding is thus consistent with that of Vining (1976), which has shown (using a non-spatial model) that if firms' technology is characterized by decreasing returns to scale, then the result would be a concave downward-sloping curve in a log–log scale instead of a straight-line, power-law relationship.

Hence, if disagglomerating effects are sufficiently large relative to the positive impact of agglomeration economies, then the power-law relationship no longer holds. The breakdown is due to the second term of Eq. (6), which specifies that greater extent of negative externalities deters agents from migrating into high-density

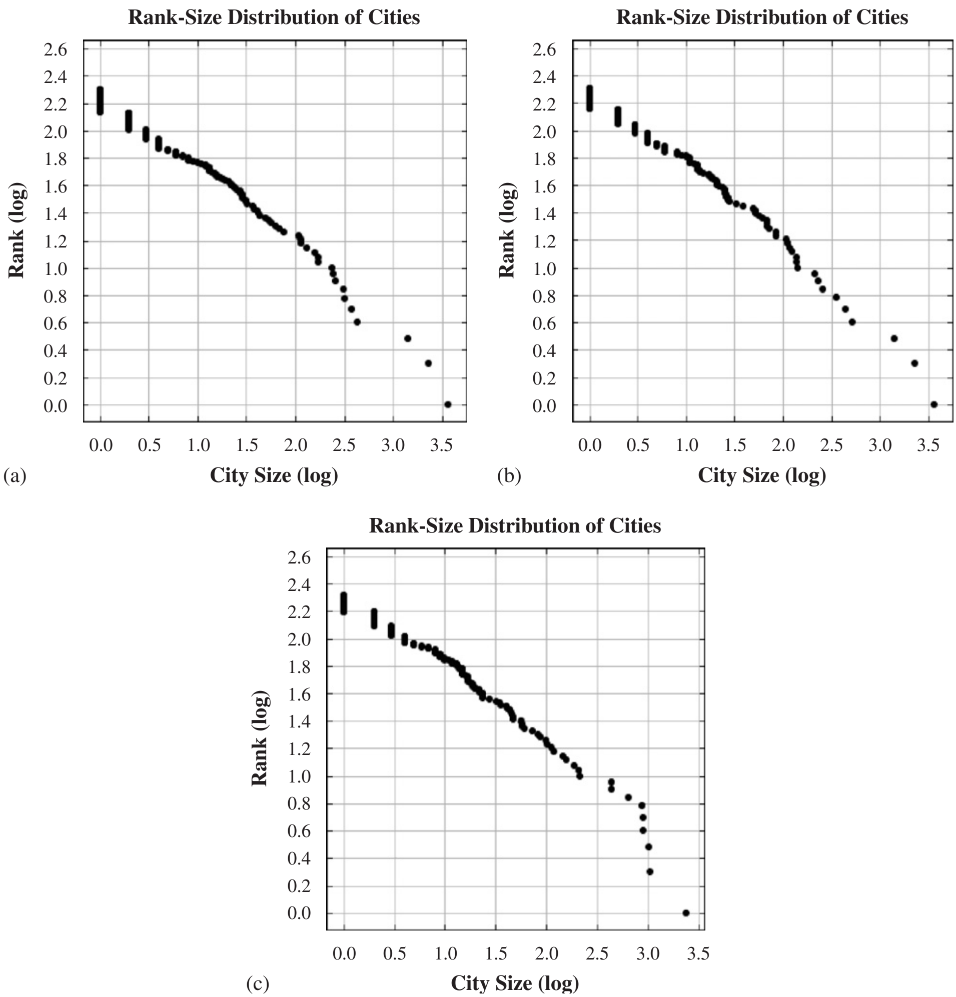

Fig. 7. The breakdown of the power law relationship when the disagglomerating force rises: (a) $c = 0.03$, (b) $c = 0.035$, and (c) $c = 0.04$. As shown, $Z = 50$, $r_{\max} = 25$, $r_{\min} = 1$, and $\phi = 0.001$. The X-axis corresponds to the population size of cities, while the Y-axis to their rank, in log–log scale.

locations that have reached the population threshold level $n_j = 1/(2c)$, thus leading to a more uniform distribution of agents across cities.

5.3. The role of space in generating the Zipf's Law

Our model shows that a power law relationship starts to emerge when agents' spatial reach exhibits sufficient variability (Fig. 2), and strengthens further as the heterogeneity of agents' spatial reach increases towards the maximum. This section examines the systematic relationship between two heterogeneous distributions: the

Table 1
Spearman rank correlation, $Corr_{n_i,r_i^{\text{avg}}}(t)$, between city size $n_i(t)$ and agents' average spatial reach $r_i^{\text{avg}}/(t)$ in the initial $t = T_0$ and in equilibrium $t = T_E$, for the "*" and the "1 15 * *" spatial reach PDFs, various grid lengths Z

PDF of spatial reach	Grid length Z	Rank correlations $Corr_{n_i,r_i^{\text{avg}}}(t)$[a]	
		Initial	Equilibrium
"*"	50	0.01	0.88
	100	0.08	0.84
	200	−0.003	0.80
	300	0.002	0.82
	400	−0.001	0.80
"1 15 * *"	50	0.05	0.93
	100	−0.004	0.79
	200	−0.003	0.70
	300	0.003	0.79
	400	−0.001	0.83

[a]Mean rank correlation over 30 simulations using a different random seed for each pair of PDF and grid length.

underlying (micro-level) agents' spatial reach and the resulting (macro-level) city-size distribution. Recall that in our model space matters because agents do not have the ability to evaluate locations that are beyond their maximum visibility reach, effectively confining agents' search to their local neighborhood. Space therefore constraints agents to travel only a limited distance at any given time, forcing long-distance migration to be carried out in several sequential periods.[11]

To start with, consider the rank correlation $Corr_{n_i,r_i^{\text{avg}}}(t)$ between the population, $n_i(t)$, and the average reach of agents, $r_i^{\text{avg}}(t) \equiv \sum_{a \in \ell_i} r_a(t)/n_i(t)$, that belong to location i for the benchmark "*" and the "1 15 * *" PDFs (Table 1).

Despite insignificant correlations $Corr_{n_i,r_i^{\text{avg}}}(t)$ in the beginning of the simulation, Table 1 suggests that the regional distribution of mobile agents experiences a significant change as the economy converges to equilibrium, which is characterized by strongly positive correlations between agents' spatial reach and city sizes.

This positive relationship between agents' mobility and city sizes highlights the role of spatial separation in generating the Zipf's Law.[12] To see this, consider the case where there are no disagglomerating forces ($c = 0$). Then for the system to stay in equilibrium, any location i must be separated from larger cities by distances at least as great as the spatial reach of location i's most mobile agents. The positive

[11]We actually allow here a small number of agents who are capable of traveling from one corner of the world to the other in a single step. Such extremely mobile agents, however, in our model are the exceptions rather than the norm.

[12]We thank a referee for pointing out the critical role of space.

correlations suggest that all mobile agents must have migrated to the larger cities that are within their spatial reach before a steady state can be attained. Otherwise, migration commences and thus such system cannot be in equilibrium. It thus would be interesting to examine empirically the prediction of our model; namely that in the long run smaller cities would have less-mobile populations on average than the larger population centers. If spatial reach can be thought of as the technological limit of agents' mobility, then Holmes and Stevens (2002) indeed found the evidence that firms with access to more sophisticated technologies tend to locate in highly-concentrated areas, which place them in a better position to exploit the productivity advantages conferred by agglomeration economies.

6. Conclusions

In this study we have proposed a spatial model of migratory autonomous agents to generate a rank-size distribution that follows the Zipf's Law as observed among US cities. Specific features of the model include explicit treatment of space, externality-driven migration, limited spatial reach, and heterogeneity of agents' mobility. We found that restrictions on the probability distribution of agents' spatial reach together with selected parameter values are sufficient for the emergence of the Zipf's Law.

The ability of our model to generate the Zipf's Law within finite times can be compared with that of Simon's and Gabaix'. In particular, the size distribution of cities in our model typically converges to an admissible equilibrium (with unit power coefficient) in less than 20 simulated periods. In contrast, Krugman (1996) reported that Simon's model is capable of generating a power coefficient that is close to one, yet can never actually reach unity. The reason has to do with a built-in property of the Simon's model, which requires a small but positive probability for newly arriving agents to form a new city instead of attaching themselves to existing cities. That property, which Krugman shows to be indispensable, prevents Simon's model to generate the exact size distribution of the Zipf's Law. In contrast, Gabaix' (1999) model requires a certain minimum city size (a 'lower reflective barrier') in order to produce the Zipf's Law. Specifically, the size of cities in the upper tail of the distribution cannot fall below a certain positive lower bound, albeit one that is allowed to grow over time. Our agent-based model requires no such restriction, and indeed some cities that initially hosted significant population shares become empty locations even in equilibria that exhibit the Zipf's Law.

In a follow-up study, we plan to extend our model to include explicitly the externality spillovers from j's neighboring locations. The significance of both physical and intangible spillovers at different levels of geographic aggregation was recently demonstrated by the empirical study of Rosenthal and Strange (2001). Thus, externality spillovers affect not only those firms sharing the same geographic area, but also neighboring firms in adjacent areas. Within the framework of our agent-based model, adding a diffusion term into Eq. (6) can capture the spillover effects from neighboring locations

$$\Psi_{j,a}(t) = n_j(t) - c \cdot n_j(t)^2 + \nabla \cdot [D_\Psi \nabla \Psi(t-1)], \tag{9}$$

where D_Ψ denotes the spillover diffusion coefficient, and $\nabla = i\partial/\partial x + j\partial/\partial y$ the spatial gradient operator. Thus for example, diffusion here can be thought of as the physical mechanism through which environmental variables (see, e.g. Fujita and Mori, 1996) transmit spillover effects.

Other extensions of this study can include explicit treatment of both the demand and supply sides of a market economy in which prices are endogenously determined. Such a general equilibrium approach in a spatial framework may explain the Zipf's Law pattern as a result of the local interactions among and between consumers and firms. In that framework, the Zipf's Law can be viewed as a complex pattern emerging from a network of many agglomeration economies; each representing an "island" of regional trading equilibrium.

Finally, our model can also be applied to understand the variability of the power coefficients outside of the US. The cross-country study of Rosen and Resnick (1980) found power coefficients that range from −0.809 (Morocco) to −1.963 (Australia) in a sample of 44 countries encompassing both developed and developing countries. Based on the 1970 census data, the estimated power coefficients are indeed reasonably close to minus unity for many of these countries with mean −1.14 in this cross-section sample. Using the more recent data for 75 countries, Soo (2005) shows that the mean is 1.11 for the full sample, which indicates the stability of the Zipf's Law over time and across countries. In the specific case of Japan, Davis and Weinstein (2002) found that not only the rank-size distribution of cities has followed Zipf's Law since 6000 years ago; in addition the identity of the cities associated with those rankings has also remained persistent.

Acknowledgements

We thank Cars Hommes, Walter Isard, Tatsuhiko Kawashima, Sidney Saltzman, and three anonymous referees for very helpful suggestions. We benefit also from the insights of seminar participants of the North American Regional Science Association International Conference in Seattle, WA, and the International Conference on Complex Systems in Nashua, NH. L.G. acknowledges the partial support of the GVOP-3.2.2-2004.07-005/3.0 (ELTE Informatics Cooperative Research and Education Center) grant of the Hungarian Government.

References

Alonso, W., 1964. Location theory. In: Friedmann, J., Alonso, W. (Eds.), Regional Development and Planning. MIT Press, Cambridge, MA.

Alvin, P., Foley, D., 1992. Decentralized, dispersed exchange without an auctioneer. Journal of Economic Behavior and Organization 18, 27–51.

Andersson, C., Hellervik, A., Lindgren, K., Hagson, A., Tornberg, J., 2003. The urban economy as a scale-free network. Physical Review E 68, 036124.
Auerbach, F., 1913. Das geset der bevölkerungskonzentration. Petermanns Geographische Mitteilungen 59, 74.
Barabasi, A.L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.
Batty, M., 2001. Editorial: cities as small worlds. Environment and Planning B 28, 637–638.
Beckman, M., 1958. City hierarchies and distribution of city size. Economic Development and Cultural Change 6, 243–248.
Christaller, W., 1933. Central Places in Southern Germany (Translated in 1966 from Die zentralen Orte in Süddeutschland by Baskin C.W.). Prentice-Hall, Englewood Cliffs, NJ.
Conlisk, J., 1996. Why bounded rationality. Journal of Economic Literature 34, 669–700.
Cullen, J.B., Levitt, S.D., 1999. Crime, urban flight, and the consequences for cities. Review of Economics and Statistics 81, 159–169.
Cyert, R., March, J.G., 1963. A Behavioral Theory of the Firm. Prentice-Hall, Englewood Cliffs, NJ.
Davis, D.R., Weinstein, D.E., 2002. Bones, bombs, and break points: the geography of economic activity. American Economic Review 92, 1269–1289.
Dobkins, L.H., Ioannides, Y.M., 2000. Evolution of size distribution: US cities. In: Huriot, J.-M., Thisse, J.-F. (Eds.), Economics of Cities. Cambridge University Press, Cambridge.
Ellison, G., Glaeser, E.L., 1997. Geographic concentration in US manufacturing industries: a dartboard approach. Journal of Political Economy 105, 889–927.
Fujita, M., Mori, T., 1996. The role of ports in the making of major cities: self-agglomeration and hub-effect. Journal of Development Economics 49, 93–120.
Fujita, M., Krugman, P., Venables, A.J., 1999. The Spatial Economy. MIT Press, Cambridge, MA.
Gabaix, X., 1999. Zipf's Law for cities: an explanation. Quarterly Journal of Economics 114, 739–767.
Gibrat, R., 1931. Les inégalités économiques. Sirey, Paris.
Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., Shleifer, A., 1992. Growth in cities. Journal of Political Economy 100, 1126–1152.
Gulyás, L., 2002. On the transition to agent-based modeling: implementation strategies from variables to agents. Social Science Computer Review 20, 389–399.
Hagerstrand, T., 1970. What about people in regional science? Papers of the Regional Science Association 24, 7–21.
Henderson, J.V., 1974. The sizes and types of cities. American Economic Review 64, 640–656.
Henderson, J.V., 1986. Efficiency of resource usage and city size. Journal of Urban Economics 19, 47–70.
Holmes, T.J., Stevens, J.J., 2002. Geographic concentration and establishment scale. Review of Economics and Statistics 84, 682–690.
Hommes, C., 2006. Heterogeneous agent models in economics and finance. In: Tesfatsion, L., Judd, K. (Eds.), Handbook of Computational Economics, Agent-based Computational Economics, vol. 2. Elsevier, Amsterdam, pp. 1109–1186.
Isard, W., 1952. A general location principle of an optimum space-economy. Econometrica 20, 406.
Johnson, N.L., Kotz, S., Balakrishnan, N., 1995. Continuous Univariate Distributions. 2nd ed., vol. 2. Wiley, New York.
Kirman, A.P., 1992. Whom or what does the representative individual represent? Journal of Economic Perspectives 6, 117–136.
Krugman, P., 1991. Increasing returns and economic geography. Journal of Political Economy 99, 483–499.
Krugman, P., 1996. The Self-Organizing Economy. Blackwell Publishers, Malden, MA.
LeBaron, B., 2002. Short-memory traders and their impact on group learning in financial markets. Proceedings of the National Academy of Sciences 99, 7201–7206.
LeBaron, B., 2006. Agent-based computational finance. In: Tesfatsion, L., Judd, K. (Eds.), Handbook of Computational Economics, Agent-based Computational Economics, vol. 2. Elsevier, Amsterdam, pp. 1187–1233.
MacLeod, W.B., 2002. Complexity, bounded rationality, and heuristic search. Contributions to Economic Analysis and Policy 1, 1–50.

Makse, H.A., Andrade, J.S., Batty, M., Havlin, S., Stanley, H.E., 1998. Modeling urban growth patterns with correlated percolation. Physical Review E 58, 7054–7062.
Manrubia, S.C., Zanette, D.H., 1998. Intermittency model for urban development. Physical Review E 58, 295–302.
Menczer, F., 2004. Evolution of document networks. Proceedings of the National Academy of Sciences 101, 5261–5265.
Mills, E.S., 1967. An aggregate model of resource allocation in a metropolitan area. American Economic Review 57, 197–210.
Mitzenmacher, M., 2004. A brief history of generative models for power law and lognormal distributions. Internet Mathematics 1, 226–251.
Murphy, K.M., Shleifer, A., Vishny, R.W., 1989. Industrialization and the big push. Journal of Political Economy 97, 1003–1026.
Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Harvard University Press, Cambridge, MA.
Newman, M.E.J., 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46, 323–351.
Page, S., 1999. On the emergence of cities. Journal of Urban Economics 45, 184–208.
Parr, J.B., 1970. Models of city size in an urban system. Papers of the Regional Science Association 25, 221–251.
Rappaport, J., Sachs, J.D., 2003. The United States as a coastal nation. Journal of Economic Growth 8, 5–46.
Roback, J., 1982. Wages, rents, and the quality of life. Journal of Political Economy 90, 1257–1278.
Rosen, K.T., Resnick, M., 1980. The size distribution of cities: an examination of the Pareto Law and primacy. Journal of Urban Economics 8, 165–186.
Rosenthal, S.S., Strange, W.C., 2001. The determinants of agglomeration. Journal of Urban Economics 50, 191–229.
Schelling, T., 1978. Micromotives and Macrobehavior. W.W. Norton, New York.
Simon, H.A., 1955. On a class of skew distribution functions. Biometrika 52, 425–440.
Simon, H.A., 1956. Rational choice and the structure of the environment. Psychological Review 63, 129–138.
Simon, H.A., Bonini, C.P., 1958. The size distribution of business firms. American Economic Review 48, 607–617.
Soo, K.T., 2005. Zipf's Law for cities: a cross country investigation. Regional Science and Urban Economics 3, 239–263.
Strogatz, S., 2005. Romanesque networks. Nature 433, 365–366.
Vining Jr., D.R., 1976. Autocorrelated growth rates and the Pareto Law: a further analysis. Journal of Political Economy 84, 369–380.
Watts, D.J., 1999. Small Worlds. Princeton University Press, Princeton, NJ.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393, 440–442.
Weber, A., 1914. Theory of the Location of Industries (Translated in 1929 by Friedrich C.J.). University of Chicago Press, Chicago, IL.
Zanette, D.H., Manrubia, S.C., 1997. Role of intermittency in urban development: a model of large scale city formation. Physical Review Letters 79, 523–526.
Zipf, G.K., 1949. Human Behavior and the Principle of Least Effort. Hafner Publishing Company, New York.