Journal of Economic Dynamics & Control 41 (2014) 38–49

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Validating an agent-based model of the Zipf's Law: A discrete Markov-chain approach

Bruno Gaujal [a], Laszlo Gulyas [b,c], Yuri Mansury [d,*], Eric Thierry [e,f]

[a] INRIA, LIG, 38330 Montbonnot, France
[b] AITIA International Inc., Budapest, H-1039, Hungary
[c] Lorand Eotvos University, Budapest, H-1117, Hungary
[d] Graduate Field of Regional Science, Department of City and Regional Planning, Cornell University, 213 West Sibley Hall, Ithaca, NY 14853, USA
[e] University of Lyon, LIP, IXXI, ENS Lyon, 69364 Lyon Cedex 07, France
[f] CNRS, LIAFA, Université Paris 7, 75205 Paris Cedex 13, France

ARTICLE INFO

Article history:
Received 1 February 2011
Received in revised form
19 January 2014
Accepted 19 January 2014
Available online 19 February 2014

JEL classification:
C61
C65
R12

Keywords:
Discrete Markov model
Agent-based models
Zipf's Law

ABSTRACT

This study discusses the validation of an agent-based model of emergent city systems with heterogeneous agents. To this end, it proposes a simplified version of the original agent-based model and subjects it to mathematical analysis. The proposed model is transformed into an analytically tractable discrete Markov model, and its city size distribution is examined. Its discrete nature allows the Markov model to be used to validate the algorithms of computational agent-based models. We show that the Markov chains lead to a power-law distribution when the ranges of migration options are randomly distributed across the agent population. We also identify sufficient conditions under which the Markov chains produce the Zipf's Law, which has never been done within a discrete framework. The conditions under which our simplified model yields the Zipf's Law are in agreement with, and thus validate, the configurations of the original heterogeneous agent-based model.

1. Introduction

The agent-based paradigm has been increasingly used to model the impact of decentralized decision-making on the observed spatial patterns (Page, 1999; Batty, 2005; Irwin, 2010). This paradigm shift has been brought about by the realization that explaining complex patterns requires explicit considerations of agent diversity (Kirman, 1992), as complex systems generally cannot be approximated by statistical averages (Page, 2011). The new paradigm, which emphasizes the role of agents that are not alike, challenges the conventional approach to modeling system behavior based on the representative agent framework. The explicit recognition of heterogeneity, however, usually comes at the cost of analytical tractability (Hommes, 2006). The present study proposes a Markov-chain platform to validate a simple agent-based model of settlement patterns. We use an analytically tractable *discrete* model to identify the core mechanism that generates the Zipf's Law in a previously developed heterogeneous agent-based model (HABM). This is the main contribution of our paper.

* Corresponding author at: Tel.: +1 607 255 4271; fax: +1 607 255 1971.
E-mail address: ysm3@cornell.edu (Y. Mansury).

http://dx.doi.org/10.1016/j.jedc.2014.02.002
0165-1889

Identifying such a mechanism is key in understanding the empirically observed distribution of people across locations. The literature on the statistical distribution that best fits the observed city size distribution is growing (see, e.g., Rosen and Resnick, 1980; Eaton and Eckstein, 1997; Dobkins and Ioannides, 2000). In this study, we calibrate our Markov model to produce a system of cities that follows the rank-size rule, known as the Zipf's Law. The rule implies that the population of the i^{th} largest city, s_i, scales according to $s_1/(r_i)^\gamma$, where s_1 is the population of the largest city, r_i is the city's ranking by size and the exponent γ is equal to unity.[1]

Whether the Zipf's Law provides the most appropriate description of the actual city size distribution has been intensely debated (see Soo, 2005; Eeckhout, 2004; Levy, 2009; Giesen et al., 2010). Indeed, across disciplines debates continue over whether the lognormal or the power law (of which the Zipf's Law is a special case) provides a better description of the observed empirical distribution (Mitzenmacher, 2004). We do not wish to engage in this controversy in this study. We take the Zipf's Law for granted and use it to anchor the size distribution produced by our Markov model. Every modeling endeavor needs an endpoint, and ours is the rank-size rule. It is, however, a flexible endpoint in the sense that the Markov-chain model is amenable to changes that would make it possible in the future to generate other distributions, such as the lognormal or the double Pareto lognormal, that compete with the Zipf's Law for best fit.

A number of theoretical studies starting with Simon (1955) have proposed random-growth models that successfully produce the Zipf's Law, including Córdoba (2008a, 2008b), Duranton (2007), Gabaix (1999), and Rossi-Hansberg and Wright (2007). In these previous studies, as in our model, the rank-size rule emerges as a stationary outcome of a stochastic process. The random-growth models also rely on the mechanism of preferential attachment (Barabasi and Albert, 1999), which, in an urban studies context, corresponds to the tendency for mobile people and activities to gravitate toward larger agglomerations (Glaeser et al., 1995).

Despite these common features, there are important differences between the above models and the HABM. Specifically, the random-growth approach treats migration as a process based on aggregate level information. Therefore, these random growth models are not directly applicable to describe the core mechanism of the HABM, where bounded rationality implies limited information access. The mechanism we discuss here is more bottom-up in the sense that changes in the size distribution are driven by the migration decision of utility-maximizing individuals. Information asymmetries are another important difference, especially because numerical simulations of our HABM suggest that it is heterogeneity in the information set that leads to the emergence of the Zipf's Law. In this paper, we strip down the rich HABM so that it is analytically tractable, and we will show that the stripped down HABM yields a closed form solution, which in turn confirms that heterogeneity is indeed the key component of the Zipf's Law.

Discrete agent-based models have been advanced as a framework that explains complex patterns at the aggregate level while maintaining consistency between the system and the individual constituents. The models are "agent based" because the main decision-making units are individuals who respond to changes in the economic environment. Spatial agent-based models have produced the Zipf's Law size distribution from the uncoordinated decisions of agents to migrate to locations that, according to their private assessment, are the most attractive (see, e.g., Axtell and Florida, 2001; Mansury and Gulyas, 2007). The discrete nature, however, combined with elements of non-linearity and non-convexity, produce spatial patterns (i.e., size distributions) that cannot possibly be derived analytically. Therefore, most agent-based models rely on numerical, computational methods.

Computer simulations have long been indispensable to urban and regional economists. However, it was Krugman (1991) who spawned a strand of literature branded as the "new economic geography" (NEG), which brought computational methods to the attention of mainstream economists (Neary, 2001). The core-periphery model Krugman (1991) proposes is not analytically tractable; simulation is the only resort. Spurred by the excitement that the NEG generated, computational methods appeared poised to become the method of choice for location theorists across disciplines (Fujita et al., 1999).

Although acceptance is growing, a search of recently published articles in the core economic journals suggests that agent-based modeling endeavors remain scanty. The key problems in applying agent-based models to spatial issues appear to be generalization and validation. The generalization problem is due to the sensitivity of numerical results on specific parameter values. For agent-based models that produce the rank-size rule, the range of parameter values where the Zipf's Law emerges and the point where it begins to unravel can only be pinned down by running more simulations. However, agent-based models often involve multiple parameters, each of which can potentially be "tuned" to generate the Zipf's Law. The vast parameter space in turn leads to a virtually infinite number of configurations. An exhaustive exploration is prohibitive, which means one cannot be certain whether the analysis is complete.

Even if it were possible to probe the entire parameter space, validating the algorithm would still be a challenge. Having implemented the basic model in a programing language, ideally the algorithm would then be validated by comparing the simulation outcomes with the prediction of the analytical model. This is not possible when the analytical model lacks a solution. Agent-based models of city systems in particular are difficult to validate because of heterogeneity, local search, positive feedbacks, and path dependencies.

Generalization and the identification of the causal mechanism are critical to agent-based modeling applications but are barely addressed in the literature, with only a few exceptions.[2] However, agent-based models that can be solved analytically are important due to the increasing recognition of computational models as a legitimate platform to examine heterogeneous

[1] Without restrictions on the power exponent, city sizes are said to be Pareto or power-law distributed.

[2] Hommes (2006) reviews agent-based models that are at least partially amenable to the analytical approach. None, however, is designed for a system of cities, although a few address spatial systems.

systems (Arthur, 2006; Tesfatsion, 2006). The present study proposes a Markov-chain model that produces the Zipf's Law as an analytical solution to an agent-based model while preserving heterogeneity and the discrete nature of contemporary migration. The mutual complementarity between analytical models and simulations has been touted elsewhere (Judd, 1997). We offer here a small vision of such integration in a study of complex spatial systems.

The rest of the paper proceeds as follows. In Section 2, we discuss our motivation for the agent-based model analyzed in this study. In Section 3, we present the corresponding general Markovian growth model and its steady-state behavior. Section 4 describes how the model can be applied to validate the simulation results of the model in Section 2. Section 5 shows how fast cities in our model converge to Zipf's Law. Section 6 concludes and discusses directions for future research.

2. Derivation of the agent-based model

Mansury and Gulyas (2007) present a heterogeneous agent-based model that was inspired by and is the extension of the "racetrack" model of Krugman (1993). The HABM is more powerful than its predecessor in at least one aspect, namely, that it generates the Zipf's Law in a system of cities, which Krugman's racetrack model is incapable of.[3]

The HABM introduced in Mansury and Gulyas (2007) makes at least four major modifications to Krugman's model:

- Agents live in a two dimensional world instead of in the one dimensional ring of Krugman's model.
- The market potential function takes a different form in our model.
- Agents have limited vision and mobility in our model, i.e., they can only evaluate the market potential of and move to locations not farther than a certain limit.
- The above limits of vision and mobility are heterogeneously distributed across agents.

To understand the main driving factors of the Zipf's Law in our model, we created a series of model versions, relaxing various combinations of the above assumptions. Numerical experiments with these model variants suggested that the key is the heterogeneity of vision (or information access), provided that there is also a limit on the number of moves an agent can make.[4] To investigate this insight further, we defined a very simple agent-based model having these characteristics.

Consider a set of n urban settlements, indexed by the subscript $i=1, 2, \ldots, n$, that are gradually populated by rural migrants who arrive every period one by one. Let $s_i(t)$ denote the population of the i^{th} urban settlement at time t. In the beginning, we assume that all locations are empty: $s_i(0)=0$ for all $1 \le i \le n$. We shall refer to non-empty locations as "cities" in the ensuing discussion. There is no reverse migration from the urban to rural areas, which implies exogenously determined and positive *net* rural–urban migration flows.[5] In addition, migrants settled in a city do not move any farther.

When rural agent a_t arrives at time t, it collects information for $1 \le k \le n$ settlements, each of which is uniformly selected, and then chooses the one with the largest population (ties are settled randomly). Here, k corresponds to the agent's limited *information access* (or vision) in terms of our previous discussion. To introduce heterogeneity, we assume that k is randomly drawn with probability α_k with domain $[1, n]$. For example, an agent whose options include all n locations has full information access to the entire system. We may also refer to k as the agents' migration reach because a greater information set corresponds to the more educated section of the population, which in the United States is known to be highly mobile,[6] even though, strictly speaking, geography plays no explicit role here.

We implemented this simple agent-based model, and our computational experiments confirmed that these assumptions are sufficient to generate the Zipf's Law in a system of (non-spatial) cities. In the remainder of the paper, we will focus on the mathematical analysis of this finding.

Our first observation is that while the main decision making rule for the migrating rural agent is optimization, i.e., choosing the location offering the largest population, the location selection procedure can be rewritten as a purely probabilistic rule. To appreciate this, let us consider the case of the (unique) location with the largest population $s_{max}(t)$. Clearly, if the arriving agent has this location in its information set, it will move to $s_{max}(t)$. On the other hand, given k, it is straightforward to calculate the probability of $s_{max}(t)$ being included in the set of sampled locations. Similarly, a location having the rth largest population will be selected if and only if it is included in a sample that contains lower ranking locations only. Note that the probability of selection does not monotonically increase in size, although it does in ranking. However, larger differences in size stabilize the ranking.

More generally, the probability that the information set of agent a_t is of size k is α_k. Because there are $\binom{n}{k}$ sets of this size and every one of them is equally likely, the probability of selecting any particular set of size k is given by

$$\alpha_k / \binom{n}{k}. \tag{1}$$

[3] For a discussion on the agent-based implementation of Krugman's original model and its behavior, see Gulyás (2002).
[4] In a spatial context, limited vision and migration reach implies that maximizing agents will eventually settle down.
[5] This net migration flow is exactly equal to 1 agent migrating from a rural to an urban area in every period.
[6] See, e.g., Dahl (2002), Greenwood (1969) and Schwartz (1976). The breakdown of mobility rates by education is available from *Geographical Mobility*, U.S. Census Bureau's *Current Population Survey*, http://www.census.gov/population/www/socdemo/migrate.html.

Let us now consider the case of the i^{th} largest location. For notational convenience, let $\sigma_t(i)$ be a permutation of the index $i=1, 2, \ldots, n$ at time t such that $s_{\sigma t(1)}(t) \geq s_{\sigma t(2)}(t) \geq \cdots \geq s_{\sigma t(n)}(t)$. That is, the index $\sigma_t(i)$ corresponds to the i^{th} largest urban place at time t. The largest city has the population $s_{\sigma t(1)}(t)$.

The i^{th} largest location, $s_{\sigma t(i)}(t)$, will be joined by agent a_t if and only if it is included in a sample that has $k-1$ other cities whose size is smaller than $s_{\sigma t(i)}(t)$. Because there are $n-i$ such cities, there are $\binom{n-i}{k-1}$ possible combinations in a sample of size k where the i^{th} largest city will be the location of choice. The maximum possible sample size is $n-i+1$ because a larger sample will necessarily include a larger location. Therefore, the number of samples that yield $s_{\sigma t(i)}(t)$ as a selection is

$$\sum_{k=1}^{n-i+1} \binom{n-i}{k-1}. \tag{2}$$

Combining (1) and (2), we obtain the probability that a_t joins the i^{th} largest location as the following:

$$p_i = \sum_{k=1}^{n-i+1} \frac{\alpha_k \binom{n-i}{k-1}}{\binom{n}{k}}. \tag{3}$$

The most important insight from Eq. (3) is that the sequence $p_1, \ldots, p_n$ is strictly positive and decreasing (i.e., "agents prefer larger cities"). We have thus expressed the agents' decision making as a path-dependent stochastic process, where the probability of an agent choosing a specific location in its information set is proportional to the location's population ranking. In particular, an agent selects the i^{th} largest location with probability p_i, and in so doing expands the location's population size so that $s_{\sigma t(i)}(t+1) = s_{\sigma t(i)}(t)+1$.

To provide microfoundations, the positive feedback effect postulated above can be integrated with a full behavioral model. Alternative sources of the positive feedback effect leading to spatial agglomeration that have been proposed in the literature include backward/forward linkages (Fujita et al., 1999), labor market pooling (Helsley and Strange, 1990), and knowledge spillovers (Glaeser et al., 1995). Such integration invariably results in models that are analytically intractable and, thus, will not be attempted here.

3. Mathematical analysis of the model

In this section, we examine the long-run distribution of the urban population of the model presented above. An alternative representation of that model is one in which each location is represented by a token, and its population size is replaced by the token's slot on the integer line. The slots are indexed according to their population size and then arranged from right to left such that the largest location is labeled "1", followed on its left by "2", and so on until the leftmost token is reached, representing the smallest location.

In this alternative representation, the arrival of a rural agent moves a token to its right by one slot. Without loss of generality, if a slot is occupied by multiple tokens, then the location at the top is selected. Similarly, when a token moves to the right joining a slot that is already occupied, then the token is inserted to the bottom of the stack. This rule implies that the relative order never changes because tokens cannot overtake one another. Let v_i denote the population size for the i^{th} ordered token obtained by sorting the tokens in descending order, $v_1(t) \geq v_2(t) \geq \cdots \geq v_n(t)$. Thus, for example, in Fig. 1, $v_1 = 9$, while $v_4 = 4 = v_5$. The state of the urban system can then be described by the set $V(t) \equiv \{v_1(t), \ldots, v_n(t)\}$. We note that $s_{\sigma_t(i)}(t) = v_i(t)$ in the distribution, but these are random variables that need not be equal in every trajectory.

Because the mechanics are identical, the original model and its token representation follow the same probability laws. In the following, we employ the token model because it is a simpler abstraction to describe how the Markov transition probabilities evolve over time. We note, however, that it is straightforward (though tedious) to derive all the results below using the original model without loss of generality. The token model is preferred because it avoids complicated combinatorics associated with the possibility for cities to overtake one another.

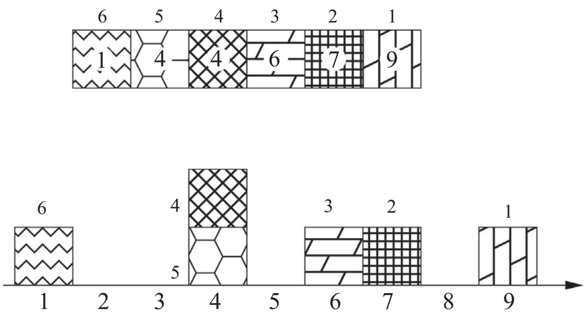

Fig. 1. Token representation of a city system's rank-size distribution. *Note*: Cities of equal size are represented by multiple tokens that are stacked on the same slot.

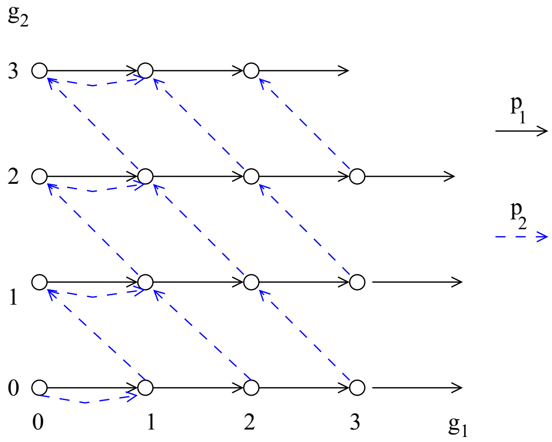

Fig. 2. The Markov chain G and its transition kernel P for $n=2$.

3.1. Markov chain analysis

The system state, $V(t) \equiv \{v_1(t), \ldots, v_n(t)\}$, corresponds to a Markov-chain process because the next period's probabilities depend only on the present state. The dynamics of the Markov process, however, cannot be directly analyzed because the transition probabilities are not time homogeneous; they depend on the current population distribution. Nonetheless, it should be clear that, over time, all tokens (populated locations) tend to shift to the right. To determine the speed of this shift, the evolution of the tokens' relative positions needs to be examined.

Our strategy is as follows. First, we show that the system eventually reaches a state in which the size differences (population gaps) between cities are strictly positive (Lemma 1). Having performed this, next we show that over time these positive population gaps become larger (Lemma 2). Then, Theorem 1 shows that the population size of the i^{th} largest city asymptotically tends to $p_i t$. As a corollary, if the probability p_i is initially Power-Law distributed, then city sizes converge to a Power Law as well. Lastly, in Section 4 we show that our stripped down HABM, introduced in Section 2, naturally leads to Power Law distributed p_i's at the beginning.

Let g_i be the size difference in terms of the population gap between settlement i and $i+1$ such that $g_i = v_i - v_{i+1}$ for $1 \le i \le n-1$ and $g_n = v_n$. Denote by h_i the index of the settlement whose token is stacked at the top of the pile above settlement i's token, i.e.,

$$h_i = \min\{j \le i, v_j = v_i\} = \min\left\{ j \le i, \prod_{k=i}^{j} g_k = 0 \right\}. \tag{4}$$

By construction, $G = (g_1, ..., g_n)$ is a Markov chain with the state space in Z^n_+, which can be grouped into 2^n subsets $\mathbf{Z}_S = \{(g_1, \ldots, g_n) \in \mathbf{Z}^n_+ \mid \forall i \in S, g_i = 0\}$, also called *faces*, for all $S \subseteq \{1, \ldots, n\}$ where the chain's kernel is space-homogeneous. The *face homogeneous* Markov chain has raised many interesting questions, but it so far lacks a general theorem to characterize its steady-state behavior (see Fayolle et al., 1995).

Definition. Let $S \subseteq \{1, \ldots, n\}$. A face $\mathbf{Z}_S = \{(g_1, \ldots, g_n) \in \mathbf{Z}^n_+ \mid \forall i \in S, g_i = 0\}$ is *homogeneous* if for any vector $v \in \mathbf{Z}^n$ and any states $G,\ G' \in \mathbf{Z}_S$the probability of reaching the state $G+v$ starting from G equals the probability of reaching $G'+v$ from G'.

The Markov chains $V = (v_1, \ldots, v_n)$ and $G = (g_1, \ldots, g_n)$ have the transition kernels O and P, respectively, such that at time $t+1$ with probability p_i for $1 \le i \le n$:

$$(v_1, \ldots, v_n) \xrightarrow{p_i} (v_1, \ldots, v_{h_i}+1, \ldots, v_n), \tag{5}$$

and

$$(g_1, \ldots, g_n) \xrightarrow{p_i} (g_1, \ldots, g_{h_i-1}-1, g_{h_i}+1, \ldots, g_n). \tag{6}$$

For example, the kernel P of the chain G is displayed in Fig. 2 for $n=2$.

Lemma 1. *The Markov chain $G = (g_1, \ldots, g_n)$ reaches a state where the population gap $g_i > 0$ for all $1 \le i \le n$ in finite time, almost surely.*

The proof involves the construction of a new Markov chain $B(t) = (b_1(t), \ldots, b_n(t))$ with a bounded state space. Appendix A shows that this new chain reaches the no-null component state with a probability of one, which means that G will also reach the state where the population gap $g_i > 0$ for all i with a probability of one.

Lemma 1 says that in our model no two cities will be of the same size in the long run. A counterexample is the case where cities are selected with equal probability, in which case the long run probability of finding two cities of the same size is non-zero.

The second step of the analysis is to show that the chain $G(t)$ drifts toward the state in which the population gaps $g_1, \ldots, g_n$ grow increasingly larger. In particular, the ratio of the gap to elapsed time will approach the difference in the probabilities of choosing the two cities associated with the gap.

Lemma 2. *After some finite time t_0, all components of the chain $G(t)$ remain positive, almost surely. Furthermore, for all i,*

$$\lim_{t\to\infty} \frac{g_i(t)}{t} = p_i - p_{i+1}. \tag{7}$$

To see the intuition behind Lemma 2, let us first construct yet another Markov chain $S(t) = (s_1(t), \ldots, s_n(t))$ with the transition kernel Q defined as follows:

(i) If X has only positive components, then

$$Q(X,Y) = \Pr(G(1) = Y | G(0) = X) = P(X,Y). \tag{8}$$

(ii) On the other hand, if X also contains null components, then,

$$Q(X,Y) = \Pr(G(\upsilon) = Y | G(0) = X) = P^{\upsilon(X)}(X,Y)$$

where the time

$$\upsilon(X) = \inf\{t \ge 0, \ \text{s.t.} G(t) \ge (1, \ldots, 1), G(0) = X\} \tag{9}$$

is almost surely finite and uniformly bounded, according to Lemma 1. Indeed, for all X containing null coordinates, let us denote by $X_2 \equiv (\max(X_1, 2), \ldots, \ \max(X_n, 2))$ the state X with all its components truncated at 2. Then,

$$\begin{aligned} \upsilon(X) &\le_{st} \inf\{t \ge 0, \ \ \text{s.t.} B(t) \ge (1, \cdots, 1), \ \ B(0) = X_2\} \\ &\le_{st} \max_{Z \in \{0,1,2\}^n} \inf\{t \ge 0, \ \ \text{s.t.} B(t) \ge (1, \ldots, 1), \ \ B(0) = Z\} \end{aligned} \tag{10}$$

which does not depend on X.

Appendix A shows that the transition kernels of S and G become one and the same as soon as $X > 0$ for all components. This means that after a finite time t_0, all of the components of G remain positive. Applying the Strong Law of Large Numbers, one then obtains Eq. (10) above.

The last statement describes how the gaps $g_1, \ldots, g_n$ grow over time. Lemma 2 can be combined with the Central Limit Theorem to determine the population gap $g_i(t)$ as $t \to \infty$. Specifically, in distribution

$$\lim_{t\to\infty} \frac{g_i(t) - t(p_i - p_{i+1})}{\sigma\sqrt{t}} = N(0, \ 1), \tag{11}$$

where $N(0, \ 1)$ is the standard normal distribution with mean 0 and variance 1, and $\sigma^2 = p_i + p_{i+1} - p_i^2 - p_{i+1}^2 + 2p_i p_{i+1}$. Note that because only one migrant arrives every period, t is also equal to the total migrants thus far, which, as $t \to \infty$, approaches the true population size. Eq. (11) standardizes the population gap and then applies the Central Limit Theorem. Now, because the gaps g_j's asymptotically become independent, and because the finite sums of independent normal distributions are also normally distributed, we can establish the (ergodic) steady-state of city sizes:

Theorem 1. *For all $i \in \{0, \ldots, n\}$, the size of the i^{th} largest city is equal to*

$$\sum_{j=1}^{n} g_j(t) = p_i t + \sqrt{t}\, S_i(t), \tag{12}$$

where $S_i(t)$ is a centered random variable that converges in distribution to a normal distribution with zero mean and bounded variance.

It is worth noting that p_i's denote the probability of selecting the i^{th} rank (i^{th} largest city independent of its identity or size). This is important because while city sizes evolve over time, the probability of selecting a given rank does not: it remains exogenous and fixed.

Corollary 1. *If the probabilities p_i's are Power-Law distributed with power coefficient γ, Eq. (12) above shows that when t is sufficiently long, then the city sizes ultimately follow a Power Law with the same power coefficient γ (because $S_i(t)/\sqrt{t} \to 0$ as $t \to \infty$).*

At first glance, Corollary 1 appears to merely push the Power Law explanation one step back. However, Section 4 below shows that the stripped down HABM, introduced in Section 2, inevitably leads to the probability p_i being Power-Law distributed, hence justifying our assumption in Corollary 1.

Remark 1. It can also be easily shown that initializing the Markov process with city sizes different from zero does not change the asymptotic behavior.

Remark 2. We have used the fact that the sequence $(p_i)_{1 \le i \le n}$ is decreasing when we apply Foster's Theorem,[7] which ensures that from a certain point in time the gap between city sizes remains strictly positive and that from that point on the i^{th}-ranked city grows almost independently from the others at rate p_i.

4. Application to the heterogeneous agent-based model

This section demonstrates how our analytical framework can be used to benchmark the agent-based computational model presented in Section 2. We will show that the Zipf's Law generated by the stripped down HABM is a special case of our Markov-chain results. For this, we need to show that the probabilities, $(p_i)_{1 \le k \le n}$, of our probabilistic agent-based model meet the conditions of our analytic results.

Recall that a decreasing sequence enables the asymptotic behavior of the face-homogenous Markov chain to be analyzed (see Remark 2). Because Eq. (3) shows that the sequence of p_i's is indeed strictly positive and decreasing, then all the results of Section 3 apply. Eq. (12) shows that in a steady-state the size of the i^{th} largest city is approximately $p_i t$. Thus, the emergence of the Zipf's Law primarily depends on the distribution of the p_i's.

Next, we show why the probability p_i is Power-Law distributed in our agent-based model, thus justifying our assumption in Corollary 1. We can obtain a closed-form solution for the sum in Eq. (3) using power series. Specifically:

$$p_i = \frac{(i-1)!(n-i)!}{n!}[x^{n-i}]\frac{S'(x)}{(1-x)^i}, \tag{13}$$

where $[x^{n-i}]$ is the coefficient of x^{n-i} in the series $S'(x)/(1-x)^i$with $S'(x)$ being the derivative of $S(x) = \sum_{k=1}^{n} \alpha_k\ x^k$.

To prove this formula, one can rewrite the sum by defining p_i in the following way:

$$\begin{aligned}\sum_{k=0}^{n-i} \alpha_{k+1} \frac{\binom{n-i}{k}}{\binom{n}{k+1}} &= \frac{(i-1)!(n-i)!}{n!}\sum_{k=0}^{n-i}(k+1)\alpha_{k+1}\frac{(n-k-1)!}{(n-i-k)!(i-1)!} \\ &= \frac{(i-1)!(n-i)!}{n!}\sum_{k=0}^{n-i}(k+1)\alpha_{k+1}\binom{n-k-1}{i-1}.\end{aligned} \tag{14}$$

The last sum has the structure of the Cauchy product of two power series. One of them is equal to the power series:

$$\Sigma_{k \ge 0}\binom{k+m}{k} x^k = \frac{1}{(1-x)^{m+1}}. \tag{15}$$

Denoting the coefficient of x^n in an $S(x)$ series as $[x^n]S(x)$, we express the sum in the following way:

$$\begin{aligned}\sum_{k=0}^{n-i}(k+1)\alpha_{k+1}\binom{n-k-1}{i-1} &= [x^{n-1}]\left(\sum_{k \ge 0}(k+1)\alpha_{k+1}x^k\right)\left(\sum_{j \ge 0}\binom{j}{i-1}x^j\right) \\ &= [x^{n-1}]\left(\sum_{k \ge 0}(k+1)\alpha_{k+1}x^k\right)\left(\sum_{j \ge 0}\binom{j+i-1}{i-1}x^{j+i-1}\right) \\ &= [x^{n-1}]x^{i-1}\underbrace{\left(\sum_{k \ge 0}(k+1)\alpha_{k+1}x^k\right)}_{S'(x)}\underbrace{\left(\sum_{j \ge 0}\binom{j+i-1}{j}x^j\right)}_{1/(1-x)^i} \\ &= \left[x^{n-i}\right]\frac{S'(x)}{(1-x)^i}.\end{aligned} \tag{16}$$

Thus, we have shown that the decision making rule of our agent-based model inevitably leads to the probability p_i being Power-Law distributed, starting from an initial distribution that differs from perfect uniformity only by a small random deviation.

Intuitively, consider the simple case of $n=2$. The full information case would result in all migrants settling in one location, the one chosen by the first migrant. In the limited information case of the HABM, however, the larger settlement would occasionally be outside of later migrants' information set. This imperfection leads to a non-degenerated equilibrium size distribution, which in turn depends on the distribution of α_k. For example, if $\alpha_1=1$ and $\alpha_2=0$, then the long-run

[7] See the proof for Lemma 2 in Appendix A.

distribution would be uniform, such that half the population settles in one city and half in the other. Now, if $\alpha_2>0$, then individuals can sometimes choose between the two cities and, in such case, will choose the larger one. Thus, $\alpha_2>0$ naturally leads to a higher concentration in one city than in the other, and the larger the $\alpha_2>0$, the larger the concentration. In between, $\alpha_1=\alpha_2=0.5$ is the intermediate case that ultimately produces the Zipf's Law.[8]

This finding based on a discrete approach complements Córdoba (2008a), who, using a stochastic continuum model, shows that the Zipf's Law obtains only if random shocks are also Power-Law distributed.

4.1. Example A: Uniformly-distributed k

We continue by presenting a simple example to demonstrate our results. We will assume that the size of the information set (k) of the rural migrant is uniformly distributed. That is, $\propto_k=(1/n)$ for all $1\le k\le n$. Under this most general condition the Markov chain produces a Power-Law, but not the Zipf's Law.

In this case, the probability p_i that the i^{th} largest city is chosen as a migration destination is equal to:

$$p_i=\frac{(i-1)!(n-i)!}{n!}[x^{n-i}]\frac{S'(x)}{(1-x)^i}, \tag{17}$$

where $[x^{n-i}]$ is the coefficient of x^{n-i} in the series $S'(x)/(1-x)^i$, with $S'(x)$ being the derivative of $S(x)=\sum_{k=1}^{n}\alpha_k\ x^k=(1/n)\sum_{k=1}^{n}x^k=(1/n)(x-x^{n+1}/1-x)$. This implies that $S'(x)=(1/n)((1-(n+1)x^n+nx^{n+1})/(1-x)^2)$. Consequently, for all $1\le i\le n$, we have:

$$\begin{aligned}
p_i&=\frac{(i-1)!(n-i)!}{n!}[x^{n-i}]\frac{1}{n}\frac{1-(n+1)x^n+nx^{n+1}}{(1-x)^{i+2}}\\
&=\frac{(i-1)!(n-i)!}{n.n!}[x^{n-i}]\frac{1}{(1-x)^{i+2}}\\
&=\frac{(i-1)!(n-i)!}{n.n!}[x^{n-i}]\frac{1}{(1-x)^{i+2}}\\
&=\frac{(i-1)!(n-i)!}{n.n!}\binom{n+1}{i+1}=\frac{n+1}{n}\frac{1}{i(i+1)}.
\end{aligned} \tag{18}$$

That is, $p_i\sim i^{-2}$, a Power Law with $\gamma=2$. Thus, because the sequence of p_i is strictly positive and decreasing, we can apply all the results of Section 3. In particular, because the p_i's follow a Power Law with $\gamma=2$ and $I(x)=\max\{i|1/p_i\le x\}\sim((n+1)x/n)^{1/2}$, the city-size distribution is a Power Law with $\gamma=1/2$.

4.2. Example B: Zipf's Law

Naturally, we are interested in conditions under which the stripped down HABM of the Section 2 model is able to produce the Zipf's Law. To this end, we study a setting where the probability of finding a rural migrant with access to k cities to choose from is decreasing in k: $\alpha_k=a_n/k$, where a_n is a normalizing constant guaranteeing that $\sum_k\alpha_k=1$. Thus, in this case, larger sets have lower probabilities of being selected, representing the more realistic case in which it is more difficult to find highly mobile agents with access of information to a large number of urban locations. We note that this construction is similar to the conditions imposed on the simulations of the original HABM (Mansury and Gulyas, 2007).

If $\alpha_k=a_n/k$, then, by definition, $S(x)=a_n\sum_{k=1}^{n}x^k/k$, and $S'(x)=a_n\sum_{k=0}^{n-1}x^{k-1}$. Therefore

$$\frac{S'(x)}{(1-x)^i}=a_n\frac{1-x^n}{(1-x)^{i+1}}. \tag{19}$$

Now, looking at the coefficients,

$$[x^{n-i}]\frac{S'(x)}{(1-x)^i}=a_n[x^{n-i}]\frac{1-x^n}{(1-x)^{i+1}}=a_n[x^{n-i}]\ \left((1-x^n)\sum_{k=1}^{n}\binom{k+i}{i}x^k\right)=a_n\binom{n}{i}. \tag{20}$$

Using Eq. (17), it follows that:

$$p_i=\frac{1}{\binom{n}{i}}\frac{a_n}{i}\binom{n}{i}=\frac{a_n}{i}. \tag{21}$$

Therefore, the probabilities p_i's are Power-Law distributed with exponent 1. Applying Theorem 1, it then follows that city sizes follow the Zipf's Law distribution.

---

[8] We thank an anonymous referee for pointing out the example.

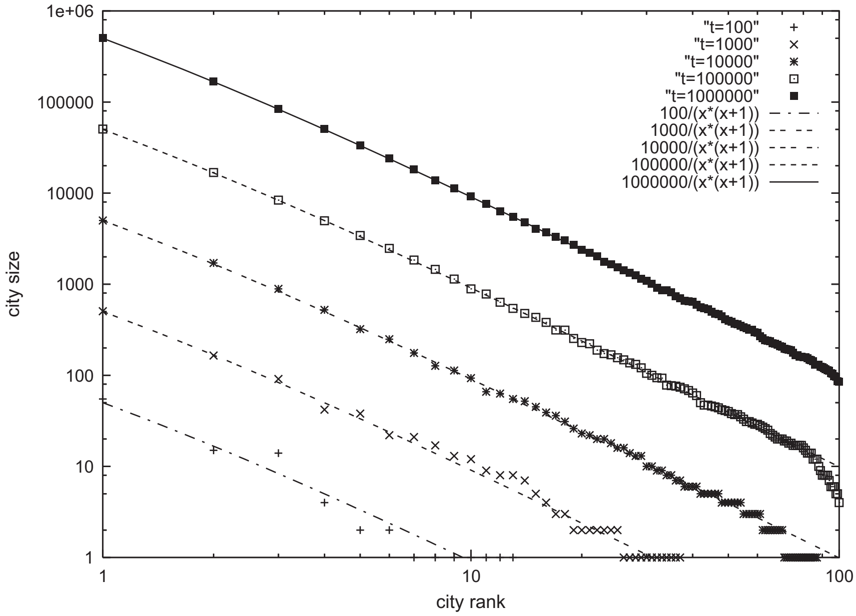

Fig. 3. Plot of city sizes against city ranks for $n = 100$, various t, in log-log scales.

5. Convergence

In Sections 3 and 4, we established our analytical results based on the steady-state behavior of the face-homogeneous Markov chain. There are, however, auxiliary functions and constants that disappear asymptotically, but it is unclear whether they influence the convergence speed in the short to medium term. In Eq. (18), for example, we know that the term $S_k(t)/\sqrt{t}$ converges to zero in probability as $t \to \infty$, yet in finite time it is difficult to determine how fast that term vanishes.

In reality, settlements converge rapidly to the Zipf's distribution, less than 150 years in the case of American cities (Dobkins and Ioannides, 2000). Thus, one may question whether our model converges to the steady state within a reasonable timeframe. The formalization of the convergence speed remains difficult, at least for now. In this section, therefore, we employ Monte Carlo simulations to demonstrate how rapidly cities in our model converge to Zipf's Law (i.e., Power Law with exponent 1). As shown in Section 4.2, Zipf's Law is obtained by setting the probability $\alpha_k = a_n/k$ (where $a_n = 1/(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n})$ is the normalizing constant).

Fig. 3 shows a typical dynamic trajectory of the rank-size distribution of cities in Monte Carlo simulations with $n = 100$ initial locations.[9] In that figure, city sizes are plotted against their rankings at time $t = 100$, 1000, 10,000, 100,000, and 1,000,000. We employed the log-log scales to highlight the resulting power-law distribution. The straight lines correspond to the theoretical rank-size plot under the asymptotic regime, namely, $p_i\, t \approx t/i$ for the i^{th}-ranked city. It can be seen that, consistent with the empirical finding of Dobkins and Ioannides (2000), the rank-size distribution of cities converges to Zipf's Law (i.e., power law with power coefficient = 1) in less than 1000 periods.

It is also noteworthy that after sufficient time has elapsed, our model reaches the point at which all cities grow at the same *scale-invariant* rate. To see this, Lemma 2 shows that at any time t, the i^{th} largest settlement (with population $s_{\sigma_{t(i)}}(t)$) grows, on average, at a rate of $1+p_i/s_{\sigma_{t(i)}}(t)$. When t is small, this rate varies across cities because it depends on the settlement's exact size and rank. However, when t is sufficiently large, the population $s_{\sigma_{t(i)}}(t)$ is approximately $p_i t$ and, therefore, settlements grow at a *mean* rate of $1+1/t$, independent of size and rank.

This scale-independent feature Gibrat (1931) originally proposed can also be found in Gabaix (1999). In contrast to Gabaix's continuous approach, however, our discrete model generates time-varying growth rates. As seen, the growth rate here decreases and tends to 1 as $t \to \infty$ due to the *additive* process that restricts a city to accommodate a maximum of only one new migrant every period.

6. Conclusions

In this paper, we provided a two-step validation of a previously developed heterogeneous agent-based model of city formation. In the quest to understand the key mechanism that leads that model to generate the Zipf's Law, we first proposed a simplified agent-based model that is also able to yield that distribution of city sizes. We then subjected this model to

[9] We used aliasing techniques in our simulation to determine which city an arriving agent will join within constant time (Walker, 1977).

mathematical analysis and derived a general result showing sufficient conditions under which the model generates Power Law distributed city sizes. This was followed by an analysis describing sufficient properties for the parameters of the simplified agent-based model to produce the Zipf's Law (a special case of the Power Law distribution). This latter result is in agreement with, and thus another validation of, the conditions proposed in Mansury and Gulyas (2007) for the original agent-based model.

The Markov-chain model we develop in this study highlights the role of random events in generating the long-run population distribution. Altering the initial probabilities in our model can lead to distinct power coefficients, including $\gamma = 1$, associated with the rank-size rule. This is an interesting feature of our model that can be deployed to explain the varying values of α in a cross section of countries (see Rosen and Resnick, 1980; Soo, 2005). While the current model is too stark to explore policy implications, its analytical results are useful for future extensions that include explicit considerations of the housing, labor, and goods markets to generate the Zipf's Law pattern from the interactions among consumers, workers, and firms.

Although both emphasize the role of random events, there are important distinctions between our model and that of Gabaix (1999). First and foremost, while agent heterogeneity in our model plays a central role in generating the Zipf's Law, it is absent in the Gabaix model. For most heterogeneous systems, a discrete, additive growth model such as ours provides a more natural platform than a continuous, multiplicative growth model such as that of Gabaix. Furthermore, beyond the obvious differences between a discrete and a continuum framework, cities in Gabaix's model, as in Córdoba (2008b), must have a minimum size (the “lower reflective barrier”) that grows continuously over time, while our model imposes no such condition to generate Zipf's Law. Lastly, cities in our model indirectly compete with one another to attract in-migrants. City sizes, therefore, must be correlated because a migrant's current choice of location precludes the possibility for that agent to reside in another city during the same period. This inter-dependence can be interpreted as competition among cities to attract young and talented migrants (Florida, 2002). In contrast, competitive interaction is absent in Gabaix (1999) because all cities exhibit independent (and identical) dynamics.

One advantage of our discrete framework is that it can be deployed to validate the algorithm of a numerical model. In particular, the analytical results can provide a benchmark to calibrate the baseline agent-based model before one pursues more complex applications for which closed-form solutions are not readily available. Our discrete framework can also be used to examine whether the results of an otherwise similar but continuous model could still hold within a discrete context, which is of interest because most contemporary urban populations usually evolve discretely, both over time and across space. For example, we can show that the comparable discrete version of Simon's continuous model does not yield a Power Law.[10]

The simplified model we used to validate the HABM has no explicit considerations for “geography”. One can interpret the parameter k (which controls the number of cities accessible to an arriving migrant) as a measure of migration reach that varies across agents. Larger values of k thus correspond to more mobile agents because these agents have access to a greater number of cities. Such an interpretation cannot be taken too far, however, because our Markov-chain model lacks spatial considerations. In the current model, an agent selects a city from a sample of size k, and which cities are included in that sample are chosen with equal probability from the entire set of n urban locations. However, geography was already represented in the original HABM as an abstract two-dimensional space.

Appendix A

A.1. Proof of Lemma 1

The proof of Lemma 1 involves the construction of a new Markov chain $B(t) = (b_1(t), \ldots, b_n(t))$ with a bounded state space. We will show first that this new chain reaches the no-null component state with a probability of one. We can show, then, that G will reach the state where the population gap $g_i > 0$ for all i with a probability of one.

Consider a new chain $B(t) = (b_1(t), \ldots, b_n(t))$ with the bounded state space $\{0, 1, 2\}^n$. Define a function f_i similar to h_i (see Eq. (4)) in construction:

$$f_i = \min \left\{ j \le i, \prod_{k=i}^{j} b_k = 0 \right\}. \tag{A1}$$

The transition kernel W of the chain B is such that for every $1 \le i \le n$:

$$(b_1, \ldots, b_n) \xrightarrow{p_i} (b_1, ..., b_{f_i - 1} - 1, b_{f_i} + 1, \ldots, b_n) \text{ if } f_i = i \text{ and } b_{f_i} < 2;$$

$$(b_1, \ldots, b_n) \xrightarrow{p_i} (b_1, \ldots, b_{f_{i-1}} - 1, b_{f_i}, \ldots, b_n) \text{ if } f_i = i \text{ and } b_{f_i} = 2;$$

$$(b_1, \ldots, b_n) \xrightarrow{p_i} (b_1, \ldots, b_{f_{i-1}} - 1, b_{f_i} + 1, \ldots, b_n) \text{ if } f_i < i.$$

[10] The proof is available upon request.

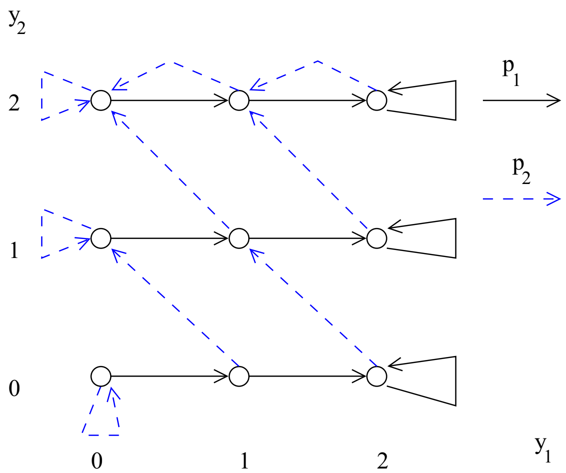

Fig. A1. The Markov chain B and its transition kernel W for $n=2$.

The finite chain B and its transition kernel W is displayed in Fig. A1. Clearly, the chain $B(t)$ reaches the state $(2, \ldots, 2)$ with positive probability.

Next, we will use induction to show that if initially $B(0) \le G(0)$ component wise, then at any time t, $B(t) \le_{st} G(t)$. Assume that $B(t) \le_{st} G(t)$. By coupling the transitions in the two Markov chains, we will show that $B(t+1) \le_{st} G(t+1)$. Specifically, with probability p_i, the two Markov chains evolve as follows:

$$G(t+1) = (g_1(t), \ldots, g_{h_i-1}(t)-1, g_{h_i}(t)+1, \ldots, g_n(t)), \tag{A2}$$

and either

$$B(t+1) = (b_1(t), \ldots, b_{f_i-1}(t)-1, b_{f_i}(t)+1, \ldots, b_n(t)) \quad \text{[Case 1]} \tag{A3}$$

or

$$B(t+1) = (b_1(t), \ldots, b_{f_i-1}(t)-1, b_{f_i}(t), \ldots, b_n(t)). \quad \text{[Case2]} \tag{A3'}$$

First, note that if $f_i = h_i$, then it is straightforward to establish that $B(t) \le_{st} G(t)$ implies $B(t+1) \le_{st} G(t+1)$ in either case. The only remaining case is when $f_i < h_i \le i$ as in Case 2 and $b_{h_i-1}(t) = 0$. Because $G(t+1)$ is non-negative by construction, this remaining case also implies that:

$$\begin{aligned}
&g_{h_i}(t+1) = g_{h_i}(t)+1 \ge b_{h_i}(t)+1 \ge b_{h_i}(t+1),\\
&g_{h_i-1}(t+1) = g_{h_i-1}(t)-1 \ge 0 = b_{h_i-1}(t) = b_{h_i-1}(t+1),\\
&g_{f_i}(t+1) = g_{f_i}(t) \ge b_{f_i}(t) = b_{f_i}(t+1) \quad (\text{if } f_i < h_i - 1),\\
&g_{f_i-1}(t+1) = g_{f_i-1}(t) \ge b_{f_i-1}(t) \ge b_{f_i-1}(t+1).
\end{aligned}$$

All of the other coordinates are unchanged in G and B with probability p_i. Thus, we have shown that $B(0) \le_{st} G(0)$ implies $B(t+1) \le_{st} G(t+1)$ for all t. To finish the proof, it is sufficient to note that because $B(t)$ reaches the state where $b_i > 0$ for all i in finite time, then so does $G(t)$ with a probability of one. This completes the proof. □

A.2. Proof of Lemma 2

We shall prove the first statement of Lemma 2. Consider the Lyapounov function min: $\mathfrak{R}^n \to \mathfrak{R}$, which has bounded average increments for the kernel Q. For all $X \in N^n$, if X contains null components, then:

$$E(\min(s_1(1), \ldots, s_n(1)) - \min(s_1(0), \ldots, s_n(0)) | (s_1(0), \ldots, s_n(0)) = X) \ge 1. \tag{A4}$$

On the other hand, if X has no null components, then:

$$E(\min(s_1(1), \ldots, s_n(1)) - \min(s_1(0), \ldots, s_n(0)) | (s_1(0), \ldots, s_n(0)) = X) \ge \min_{i=1}^{n}(p_i - p_{i+1}). \tag{A5}$$

Using an adequate version of Foster's Theorem (see, e.g., Lamperti, 1960), it can be shown that the chain S is transient and that $\min(s_1(t), \ldots, s_n(t))$ approaches infinity as $t \to \infty$. Thus, there exists a finite time τ such that for all $t \ge \tau$, $s_i(t)$ remains strictly positive for all i. Furthermore, the transition kernels of S and G coincide as soon as $X > 0$ for all components. Therefore, there exists two finite random variables n_0 and m_0 such that $S(t) = G(t+m_0)$ for all $t \ge n_0$. This means that after a finite time t_0, all of the components of G remain positive. This ends the proof for Lemma 2's first statement.

For the second statement of Lemma 2, consider the i.i.d. sequence $(\delta_j)_{j \in N}$ such that:

$\delta_j = 1$ with probability p_i;
$\delta_j = -1$ with probability p_{i+1}; and
$\delta_j = 0$ with probability $1 - p_i - p_{i+1}$.

Because $g_i(t) > 0$ almost surely after a finite time t_0 according to the first statement, and $g_i(t) - g_i(t_0) = \sum_{j=t_0}^{t} \delta_j$ for all $t \geq t_0$, the Strong Law of Large Numbers yields:

$$\lim_{t \to \infty} \frac{g_i(t)}{t} = \lim_{t \to \infty} \frac{\sum_{j=t_0}^{t} \delta_j}{t} = p_i - p_{i+1} \text{ almost surely.} \tag{A6}$$

This completes the entire proof for Lemma 2. □

References

Arthur, W.B., 2006. Out-of-equilibrium economics and agent-based modeling. In: Tesfatsion, L., Judd, K. (Eds.), Handbook of Computational Economics, vol. 2: Agent-Based Computational Economics, Elsevier, Amsterdam, pp. 1551–1564. (Chapter 32).
Axtell, R., Florida, R., 2001. Emergent cities: a microeconomic explanation for Zipf's Law. Comput. Econ. Finance, 154.
Barabasi, A.L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.
Batty, M., 2005. Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals. MIT Press, Cambridge, MA.
Córdoba, J. -C., 2008a. On the distribution of city sizes. J. Urban Econ. 63, 177–197.
Córdoba, J. –C., 2008b. A generalized Gibrat's Law. Int. Econ. Rev. 49, 1463–1468.
Dahl, G.B., 2002. Mobility and the return to education: testing a Roy model with multiple markets. Econometrica 70, 2367–2420.
Dobkins, L.H., Ioannides, Y.M., 2000. Evolution of size distribution: U.S. Cities. In: Huriot, J.M., Thisse, J.F. (Eds.), Economics of Cities, Cambridge University Press, Cambridge, UK.
Duranton, G., 2007. Urban evolutions: the fast, the slow, and the still. Am. Econ. Rev. 97 (1), 197–221.
Eaton, J., Eckstein, Z., 1997. Cities and growth: theory and evidence from France and Japan. Regional Sci. Urban Econ. 27, 443–474.
Eeckhout, J., 2004. Gibrat's law for (all) cities. Am. Econ. Rev. 94, 1429–1451.
Fayolle, G., Malyshev, V.A., Menshikov, M.V., 1995. Topics in the Constructive Theory of Countable Markov Chains. Cambridge University Press, Cambridge, UK.
Florida, R., 2002. The Rise of the Creative Class. Basic Books, Cambridge, MA.
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. Q. J. Econ. 114, 739–767.
Gibrat, R., 1931. Les Inégalités Économiques. Sirey, Paris.
Giesen, K., Zimmermann, A., Suedekum, J., 2010. The size distribution across all cities – double Pareto lognormal strikes. J. Urban Econ. 68 (2), 129–137.
Glaeser, E.L., Scheinkman, J.A., Shleifer, A., 1995. Economic growth in a cross-section of cities. J. Monet. Econ. 36, 117–143.
Greenwood, M.J., 1969. An analysis of the determinants of geographic labor mobility in the United States. Rev. Econ. Stat. 51, 189–194.
Gulyás, L., 2002. On the transition to agent-based modeling: implementation strategies from variables to agents. Soc. Sci. Comput. Rev. 20, 389–399.
Helsley, R.W., Strange, W.C., 1990. Agglomeration economies and matching in a system of cities. Reg. Sci. Urban Econ. 20, 189–212.
Hommes, C., 2006. Heterogeneous agent models in economics and finance. In: Tesfatsion, L., Judd, K. (Eds.), Handbook of Computational Economics, vol. 2. , Elsevier, Amsterdam, pp. 1109–1186. (Agent-Based Computational Economics).
Irwin, E., 2010. New directions for urban economic models of land use change: incorporating spatial dynamics and heterogeneity. J. Reg. Sci. 50 (1), 65–91.
Judd, K.L., 1997. Computational economics and economic theory: substitutes or complements? J. Econ. Dyn. Control 21 (6), 907–942.
Kirman, A.P., 1992. Whom or what does the representative individual represent? J. Econ. Perspect. 6, 117–136.
Krugman, P., 1991. Increasing returns and economic geography. J. Polit. Econ. 99, 483–499.
Krugman, P., 1993. On the number and location of cities. Eur. Econ. Rev. 37, 293–298.
Lamperti, J., 1960. Criteria for the recurrence or transience of stochastic process I. J. Math. Anal. Appl. 1, 314–330.
Levy, M., 2009. Gibrat's Law for (all) cities: comment. Am. Econ. Rev. 99, 1672–1675.
Mansury, Y., Gulyas, L., 2007. The emergence of Zipf's Law in a system of cities: an agent-based simulation approach. J. Econ. Dyn. Control 31, 2438–2460.
Mitzenmacher, M., 2004. A brief history of generative models for power law and lognormal distributions. Internet Math. 1, 226–251.
Neary, J.P., 2001. Of hype and hyperbolas: introducing the new economic geography. J. Econ. Lit. 39 (2), 536–561.
Page, S., 1999. On the emergence of cities. J. Urban Econ. 45, 184–208.
Page, S., 2011. Diversity and Complexity. Princeton University Press.
Rosen, K.T., Resnick, M., 1980. The size distribution of cities: an examination of the Pareto Law and primacy. J. Urban Econ. 8, 165–186.
Rossi-Hansberg, E., Wright, M.L.J., 2007. Urban structure and growth. Rev. Econ. Stud. 74 (2), 597–624.
Schwartz, A., 1976. Migration, age, and education. J. Polit. Econ. 84, 701–720.
Simon, H.A., 1955. On a class of skew distribution functions. Biometrika 52, 425–440.
Soo, K.T., 2005. Zipf's Law for cities: a cross country investigation. Reg. Sci. Urban Econ. 3, 239–263.
Tesfatsion, L., 2006. Agent-based computational economics: a constructive approach to economic theory. In: Tesfatsion, L., Judd, K.L. (Eds.), Handbook of Computational Economics, vol. 2. , North-Holland, Amsterdam, pp. 831–880.
Walker, A.J., 1977. An efficient method for generating random variables with general distributions. ACM Trans. Math. Softw. 3, 253–256.