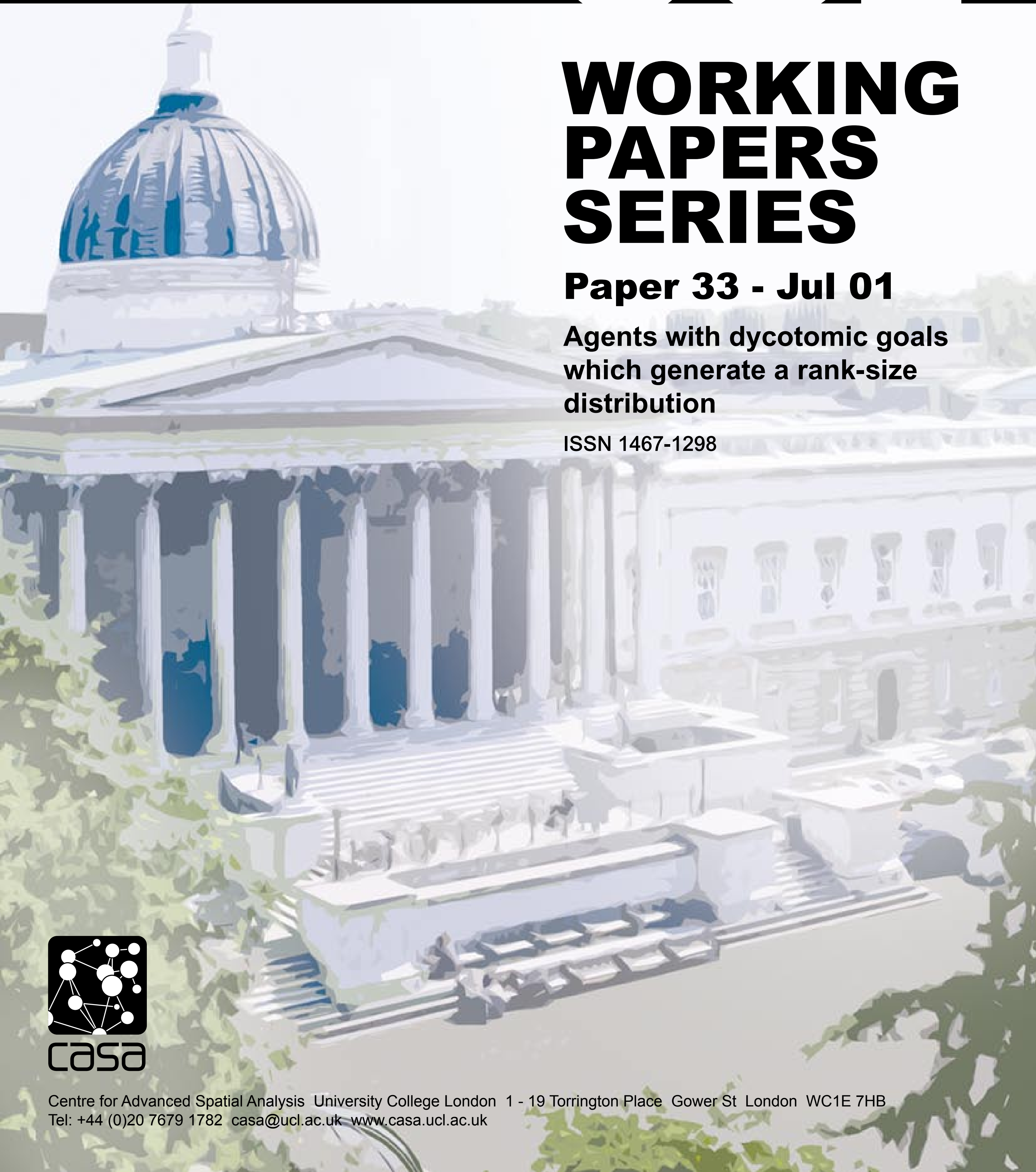
UCL CENTRE FOR ADVANCED SPATIAL ANALYSIS



WORKING PAPERS SERIES

Paper 33 - Jul 01

Agents with dycotomic goals which generate a rank-size distribution

ISSN 1467-1298

Centre for Advanced Spatial Analysis University College London 1 - 19 Torrington Place Gower St London WC1E 7HB
Tel: +44 (0)20 7679 1782 casa@ucl.ac.uk www.casa.ucl.ac.uk

Centre for Advanced Spatial Analysis
University College London
1-19 Torrington Place
Gower Street
London WC1E 6BT

[t] +44 (0) 20 7679 1782
[f] +44 (0) 20 7813 2843
[e] casa@ucl.ac.uk
[w] www.casa.ucl.ac.uk

AGENTS WITH DYCOTOMIC GOALS WHICH GENERATE A RANK-SIZE DISTRIBUTION
Ferdinando Semboloni

http//www.casa.ucl.ac.uk/paper33.pdf

Date: July 2001

ISSN: 1467-1298

Agents with dycotomic goals which generate a rank-size distribution

Ferdinando Semboloni
University of Florence
e-mail:semboloni@urba.arch.unifi.it

23rd July 2001

1 Introduction

Many explanations have been proposed for the rank-size rule or power law in city size distribution based on a probabilistic process [4]. These explanations are usually opposed to that proposed by Zipf [11] who explained the rank-size rule as the result of the application of the principle of least effort. In his opinion, by using this principle, it is possible to find an equilibrium between the two opposite forces of diversification and of unification. In fact, because the main components of the system are resources, people and products, the first force brings people near to resources, and the latter brings products near to people. Even these notions are simple, and are accepted in the spatial economic field [5] it is not clear how a rank-size rule can be derived from it[2].

In this paper I will show how a rank-size distribution can be generated by using multi-agent interaction which uses a probabilistic law to obtain opposing goals that correspond to unification and diversification forces. This paper is divided in two sections: the first section presents a model based on agents pursuing opposite goals; the second discusses the model in relation to the previously proposed models.

2 The model

The rank-size rule states that a set of events, when ranked by frequency or size shows the following property:

$$r_i^{\alpha} S_i = K \tag{1}$$

where S_i is the frequency or size of the i event, r_i is the rank, the most frequent event having rank equal one, and $\alpha \sim 1$ is an exponent.

In the present model I consider a set of cities in which resources are equally distributed: for instance one unit of resource in each city. More cities signify more available resources for the entire population. Agents live and work in each city; they represent groups of inhabitants, both as producers and as consumers and are supposed to have two goals: to produce and sell goods and to utilize resources. In the first case, because an agent must sell the good, he/she prefers to live near the greatest number of consumers. i.e. in a big city; in the second case an agent prefers to live where resources are shared among the minimum number of people. If all agents pursue the first goal the result is a single big city, i.e. the city which is ranked one, and total utilized resources will be one unit; however if the agents behave in the second way, the result is N cities of size equals one and total utilized resources will be N units. Thus in the first condition, the agent would live in first rank cities, or with a minimum rank, while in the second condition, agents would live in a city having a minimum size. First of all let us suppose that the differences among ranks and sizes are perceived in a way that the distances between things near by are emphasized in relation to those further away. This result is produced by a logarithmic transformation: in the first case an agent would like to minimize $\log(r_i)$ while the second would minimize $\log(S_i)$. Nonetheless, if agents in one period precisely minimize the function of rank and in the other period the function of size, then the result will be one big city and a lot of little ones of the same size which are equal to one. In order to obtain a result similar to that observed, the behavior of the agents needs to be disturbed by an event brought about by the different preferences among agents, i.e., limited knowledge, etc. The result is that in the first case agents will locate in the city i where the score:

$$A_i = \left[\log\left(\frac{r_i}{N}\right) + \lambda(W)\right] \tag{2}$$

has the minimum value. In this expression λ, a stochastic disturbance term, is a random Gaussian variable with a mean equal to zero, and a standard deviation equal to 1, and W is a weight related to the disturbance in estimation of the score. The rank r_i is divided by N, the total number of cities with $S_i > 0$, in order that the result will range [0-1].

In the second case agents first choose the desired size S_k and then randomly among the cities as further explained. The desired size is chosen among all the available sizes, ranging from 1, which is the minimum size of a city, to S_{max}, and having the minimum score:

$$B_k = \left[\log\left(\frac{S_k}{S_{max}}\right) + \lambda(W)\right] \tag{3}$$

In the previous expression the desired size is divided by the maximum so that the result will range [0-1]. In order to live in a city of the desired size S_k, the agent chooses at random a city i having $S_i = S_k - 1$. In fact with the location of the agent the size will automatically increase by one. This aspect is interesting in the lowest level cities because it induces the growth of a new city, or better stated, the activation of an existing site. In fact the city may be abandoned by the last agent and then be reactivated by a different agent which decides to live in the previously abandoned city.

Because an agent's behavior is determined by economic convenience, the agent will look for a minimum-sized city only if the total resources utilized are lower than the number requested. Otherwise an agent will look for the minimum rank in order to cluster with other agents. The total requested resources depend by the total population and by the technology which establishes the quantity of resources in relation to population. Because the quantity of resources is proportional to the number N of cities, the demand of resources is expressed in number of cities and is a function of the total number of agents S as in the following equation:

$$N = S^{\gamma} \tag{4}$$

where γ is a parameter which is related to the available technology. In fact the lower the parameter the lower the needed quantity of resources per inhabitant.

The model has been applied to a set of 500 cities with a number of agents equal to 1000 assigned at random to the cities. At each iteration an agent is chosen at random and is relocated in one of the cities by using equation 2 if the number of existing cities is greater than desired (equation 4) or 3 otherwise. After the relocation of the agent the size, as well as the rank of cities are re-calculated.

The two crucial parameters are γ and W, the parameter controlling the disturbance term. First of all a value for γ has been established in a way that the resulting value for N is consistent with $\alpha = 1$ in equation 1. This is obtained when $\gamma \sim 0.76$. By using this value for γ the value of parameter W has been estimated by minimizing the correlation coefficient related to the linear regression in double logarithmic coordinates of the rank-size distribution (equation 1). As the figure 1 shows the correlation coefficient is minimum when $W \sim 1.7$. By using $\gamma = 0.76$ and $W = 1.7$ as the central values the plots of the resulting rank-size distribution for nine combinations of the values of these parameters are shown in figure 2. Parameter γ affects the slope of the graph, i.e., the hierarchical character of the system, while W affects the concavity or the convexity of the graph. Because W is a measure of the disturbance it can be related to the quantity of diverse factors influencing the size distribution. If the number of these factors is reduced

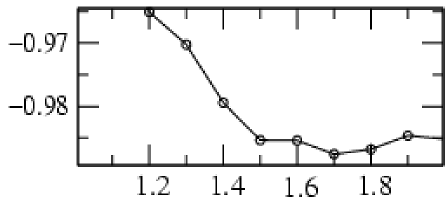


Figure 1: X axis: W, Y axis: correlation coefficient

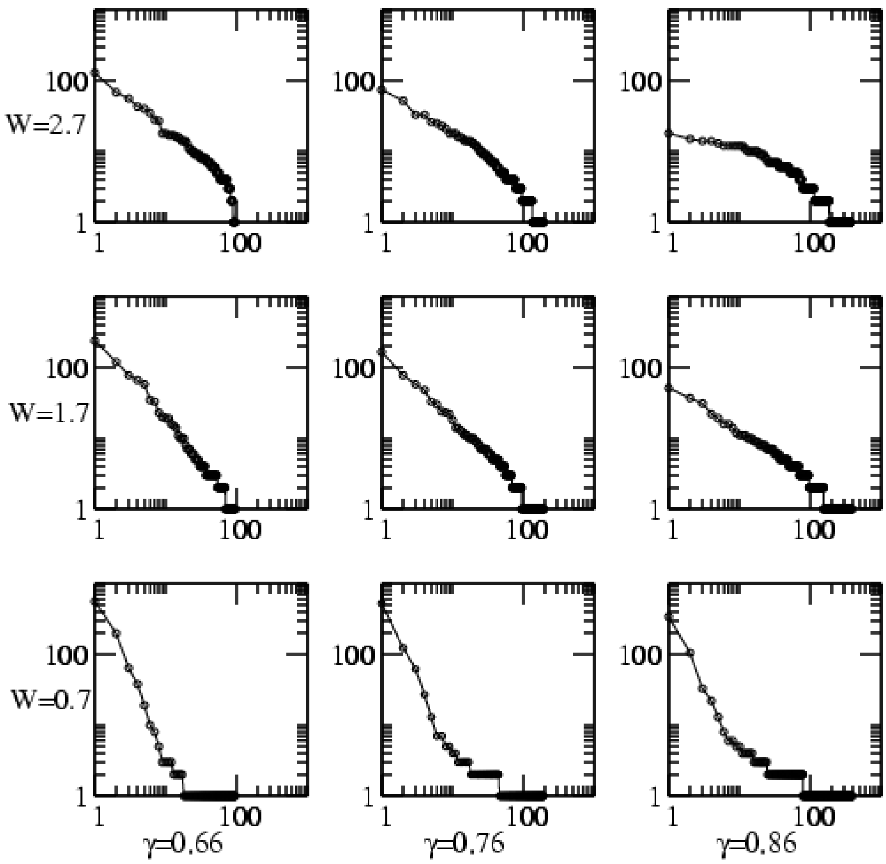


Figure 2: Rank-size graphs (X axis: rank, Y axis: size) obtained by varying γ and W. Double logarithmic scale

[1] and the value for W is low, then a distribution similar to that of the primate city is obtained; in turn, when the value of W is higher then the number of factors affecting the choice is bigger, and the differences among ranks or sizes are less important.

The convergence of the system to the rank-size distribution has been evaluated by considering the slope of the distribution and the correlation coefficient as previously stated. As figure 3 shows, the system became stable after about 3,000 iterations. If

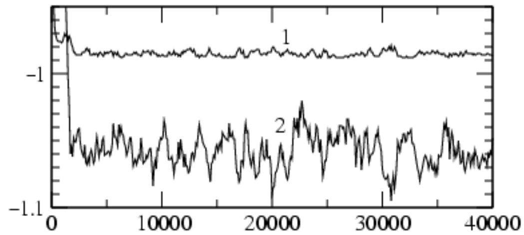


Figure 3: X axis: iteration. Y axis, 1: correlation coefficient, 2: estimated value for α (equation 1)

the distribution is stable, the rank of a city can be changed considerably. In order to evaluate this aspect, calculations have been made for an index:

$$I(t + \Delta t) = \frac{\sum_i \mid r_i(t) - r_i(t + \Delta t) \mid [S_i(t) + S_i(t + \Delta t)]}{N \sum_i [S_i(t) + S_i(t + \Delta t)]} \tag{5}$$

This index considers the variation in rank weighted with the size of the city. The resulting plot is shown in figure 4, where $\Delta t = 100$. The index grows in the first phase of the simulation and then becomes quite stable around value 0.3.

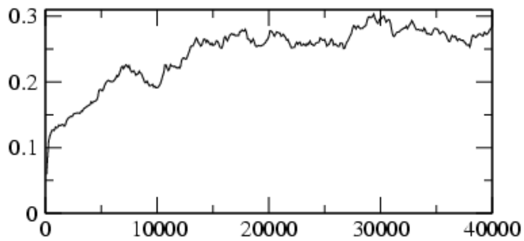


Figure 4: X axis: iteration, Y axis: index I

Even if the model does not considers growth it can be included. It is sufficient that an agent could be generated by an other agent. Initially the generated agent is located in the city where the generating agent is located. Figure 5 shows the resulting distribution by supposing a probability 0.1 to 0.3 that a new agent will be generated during one iteration.

3 Discussion

In order to discuss the proposed model, let us compare it with a simple Monte Carlo method for the generation of a rank-size distribution. In fact, let us suppose that the same population of agents is located at random on a set of cities. At each iteration an agent is chosen at random and is relocated in the city having the smallest score calculated as rank multiplied by the size as in the following expression:

$$X_i = r_i^{\alpha} S_i \tag{6}$$

After the relocation of the agent the size, as well as the rank of cities are re-calculated. This method is nothing but a Monte Carlo generation of a function and is presented here in order to better understand the functioning of the proposed model. In fact, if we consider a logarithmic transformation of the previous equation we obtain:

$$\log(X_i) = \alpha \log(r_i) + \log(S_i) \tag{7}$$

It is interesting to consider that the agent of the present model at different periods tries to minimize one of the two parts of the previous expression.

In order to understand in depth the functioning of the methods, equation 2 has been iteratively applied to a set of ranked cities. The resulting probability density distribution in dependence of various values for W is shown in figure 6. This distribution is similar to that resulting from the generalized rank-size function proposed by Mandelbrot[7]. The utilization of the logarithm is crucial. In fact when the equation 2 is applied avoiding the logarithmic transformation, the result is an exponential function (see figure 6). In essence the functioning of the present model is similar to that of Mandelbrot[7][6]. In fact, a similarity can be drawn between the way in which a random process produces words and that of choosing a city in which to live when the equation 2 is applied. In this case an agent can be supposed to choose the city in which to live beginning from the first rank. With an established probability the agent will consider the next city in the rank or decide to choose the current city. This method is similar to the random generation of words, where with an established probability a character is added to the previous characters or a space is chosen and the word is terminated. In the Mandelbrot model the changing of variable from length to the rank of the different words of the same length, is able to generate a rank-size distribution which, in turn, in the present model depend on the utilization of logarithm. In fact the logarithmic transformation correlates the probability of choosing the current city with city rank, i.e., with the size of the city.

The Simon model[9] , differs from the Mandelbrot given that the hierarchy is practically established exogenously through the growth process, yet it is similar because the city to be lived in is chosen with a probability which is proportional to the existing

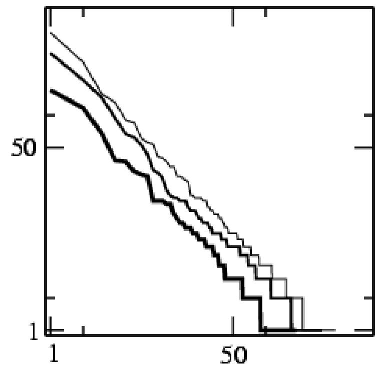


Figure 5: The rank size distribution (X axis: rank, Y axis: size) obtained by the application of 0.1, 0.2, and 0.3 grow rate. Logarithmic scale

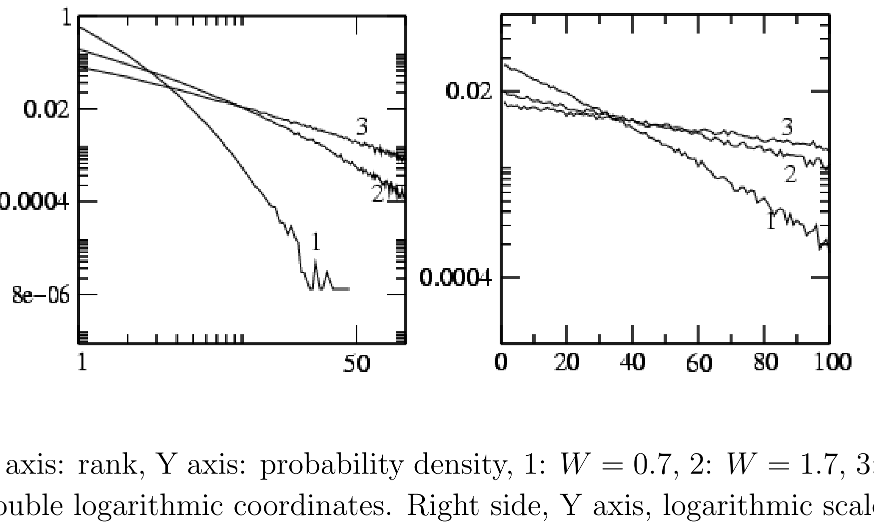


Figure 6: X axis: rank, Y axis: probability density, 1: $W = 0.7$, 2: $W = 1.7$, 3: $W = 2.7$ Left side, double logarithmic coordinates. Right side, Y axis, logarithmic scale

population. By using this method the relative distribution of population would remain unchanged. The growth process is responsible for increasing the probability that the first ranked cities will be chosen as a living place. Anyway the similarity of these two models lies in the fact that they consider the rank-size hierarchy as the result of a process oriented to the most sized event and that the lower tail results only by the diminishing of that process. The presented model is different because there is a possibility to prefer the first ranked cities or the lower level city because resources are considered as important as the clustering. These two preferences interact because size and rank may change after the relocation of an agent thus affecting all the following dynamics.

Among the probabilistic model utilized for the generation of rank-size distribution, those based on proportional random growth [4] [3] seem to implicitly include opposite forces. In fact this random growth can be realized with a random growth rate with a fixed mean and standard deviation or, as the case of Manrubia and Zanette[10], as an intermittent growth rate taking at random the values 0 or 2. The Marsili and Zhang [8] model presents another version of the model based on individuals. In all these models the agglomeration results from the temporal cluster of positive growth while dispersion or decreasing of growth is produced by the temporal cluster of negative growth. In fact the dynamic of these models is based on local growth and interaction is limited to the established total number of inhabitants and a diffusion process which is needed in order to avoid a population fall to zero in these cities. The present model is different from these for two inter-connected aspects. First of all it is clearly based on agents, second the interaction is global. In fact these agents are able to compare all the opportunities offered by the cities which are considered as a collection of agents and not capable of autonomous growth.

4 Conclusion

The model, which has been presented, is based on agents with dycotomic goals. This model explains how the two Zipfian forces are able to produce a rank-size distribution. I will be delighted if someone would send critique or comment to my e-mail address.

5 Acknowledgment

This study was carried out during my one-month stay at CASA and was later revised in Florence. While in London, I discussed an earlier version of the the study with Michael Batty, Narushige Shiode, and Paul Torrens, and in Florence with Tito Arecchi, Antonio Politi and Stefano Ruffo. All provided stimulating criticism and offered many suggestions. I am indebted to all. Obviously all errors, mistakes and the responsibility for the general approach, etc. are all mine. This paper is a part of a research project, funded by MURST (Italian Ministry of University and Research), concerning "Knowledge Engineering in Planning Process". The project is co-ordinated by Prof. Lidia Diappi, Polytechnic University of Milan.

References

[1] B. J. L. Berry. City size distributions and economic developmen. *Economic Development and Cultural Change*, 9:593–407, 1961.

[2] B. J. L. Berry and W. L. Garrison. Alternate explanations of urban rank size relationships. *Annals of the Association of American Geographers*, 48:83–91, 1958.

[3] A. Blank and S. Solomon. Power law and cities population. *cond-math*, 0003240v2, 2000.

[4] X. Gabaix. Zipf 's law for cities: An explanation. *Quarterly Journal of Economics*, CXIV:739, 1999.

[5] P. R. Krugman. *The Self-Organizing Economy.* Blackwell Publishers, Oxford, 1996.

[6] W. Li. Random texts exhibit zipf's-law-like word frequency distribution. *IEEE Transactions on Information Theory*, 6:1842–1845, 1992.

[7] B. Mandelbrot. An informational theory of the statistical structure of language. In W. Jackson, editor, *Symp. Applied Communications Theory*, pages 486–500. Betterworth, 1953.

[8] M. Marsili and Y.-C. Zhang. Interacting individuals leading to zipf 's law. *Physical Review Letters*, 80(12):2741–2744, 1999.

[9] H. Simon. On a class of skew distribution functions. *Biometrika*, 44:425–440, 1955.

[10] D. H. Zanette and S. C. Manrubia. Role of intermittency in urban development: A model of large-scale city formation. *Physical Review Letters*, 79:523–526, 1997.

[11] G. Zipf. *Human Behavior and the Principle of Least Effort.* Addison Wesley, Cambridge, MA., 1949.