

**行政院國家科學委員會專題研究計畫成果報告**

**「階層大小法則」之理論探討與動態過程之模擬研究**

**The Theoretical Derivation and the Dynamic Process Simulation of  
the Rank Size Rule**

計畫類別：個別型計畫

計畫編號：NSC 89-2415-H-004-041

執行期間：89 年 8 月 1 日至 91 年 1 月 31 日

計畫主持人：陳心蘋

執行單位：國立政治大學經濟系

中華民國 90 年 1 月 24 日

中文摘要

由於實證資料顯示瑞普夫定理 (Zipf's law) 對不同經濟結構與不同時代的國家中都市大小的分佈有共同的解釋能力，此定理已被界定為都市成長模型的基本條件。吉伯特定理 (Gibrat's law) 被證明為瑞普夫定理的充份條件。它是以純統計的方法來解釋都市的階層大小法則 (Rank-Size Rule)。本文的目的在探討一般的經濟模型與純統計觀點的吉伯特定理之間的相關性。分析結果顯示一般的經濟模型，在地區選擇機率是 logit 型態的假設下可推導出區域都市人口的成長率有共同的平均值與變異數。根據吉伯特定理，這個同質分配的都市成長率特性隱含了區域都市的極限分配會趨近於瑞普夫定理。在沒有 logit 型式的選擇機率假設下，都市極限分配的型態會受到區域居民的外佈效果和廠商的聚集經濟特性的影響。當總外佈效果趨近一共同的臨界值的情形下，都市的極限分配會趨近於瑞普夫型態。

實證資料顯示冪次法則 (Power law) 是許多俱有自我組織結構 (self-organized structure) 的複雜系統 (complex system) 的共同特性；而區域科學中解釋城市大小分佈的普瑞夫定理 (Zipf's law) 是符合冪次法則資料中最受到頻繁討論的分佈。本文主要是以動態複雜系統中自我組織結構的角度來探討冪次法則與複雜系統的相關性以及可能影響冪次參數值的原因。並進而以動態複雜系統的特性檢視以吉伯特定理 (Gibrat's law) 解釋普瑞夫定理 (Zipf's law) 的充分性。研究結果顯示吉伯特定理所指出的同質的隨機成長率僅能保證冪次法則的實現；而隨時間漸緩遞減標準差的隨機成長率才能衍生出趨近瑞普夫分配的城市分佈型態。隨時間漸緩遞減標準差的隨機成長率是來自於都市成長的複雜系統中逐漸增進的潛在聯繫與階層組織相互影響的敏感度。城市間的潛在聯繫與相互影響敏感度越多越高，城市間的區位利益差異越小，因此都市的成長率的差異會越小。系統中潛在聯繫與敏感度增進的速度會影響冪次分佈的冪次參數值與分佈收斂的速度。

關鍵詞：瑞普夫定理，吉伯特定理，階層大小法則，自我組織臨界性，複雜系統

# The Theoretical Derivation and the Dynamic Process Simulation of the Rank-Size Rule

## Abstract

Zipf's law has been considered the minimum criterion for the city growth model due to its robust empirical evidence across various types of countries and dates. Gibrat's law is proved to be the condition for the emergence of Zipf's limit distribution by a statistical method. This paper shows that a general economic growth model with logit choice probability will converge to Zipf's law in the steady state. The growth process derived from the gravity type optimal migrant flow is homogeneous. This implies that the gravity type optimal migrant flow has limit distribution converging to Zipf's pattern based on Gibrat's law. Without the assumption of discrete choice probability, the pattern of region's limit distribution is closely related to the property of the externality effects of residents and firms. Zipf's pattern will emerge as the limit distribution if the externalities from resident and firm converge to an asymptotic value as city size getting large.

Power law has been shown to be a common feature of many self-organized complex systems, and Zipf's law in regional science is the most famous of all these distributions. This paper shows that the assumption of homogeneity of the random growth process as assumed in Gibrat's law will generate city size distribution as power law. However, Gibrat's law does not necessarily generate Zipf's limiting pattern. City distribution could possibly converge to a Zipf's pattern limiting distribution only with a diminishing decreasing standard deviation of the random growth rate. Moreover, the value of the diminishing rate of the standard deviation of city growth rate determines the speed of the convergence and the value of the converged slope. The homogeneous random evolving process is the essential underlying feature, which generates the common power law property of many complex systems. Nevertheless, the variation of the changing rate of increased potential connections and the sensitivity of interactions among cities are the major reasons for the differences of the slopes among self-organized systems.

**Keywords:** Zipf's Law, Gibrat's Law, Rank-Size Rule, Self-organized criticality, Complex systems

# **The Theoretical Derivation and the Dynamic Process Simulation of the Rank Size Rule - Gibrat's Law and the Growth of Cities**

## **Abstract**

Zipf's law has been considered the minimum criterion for the city growth model due to its robust empirical evidence across various types of countries and dates. Gibrat's law is proved to be the condition for the emergence of Zipf's limit distribution by a statistical method. This paper shows that a general economic growth model with logit choice probability will converge to Zipf's law in the steady state. The growth process derived from the gravity type optimal migrant flow is homogeneous. This implies that the gravity type optimal migrant flow has limit distribution converging to Zipf's pattern based on Gibrat's law. Without the assumption of discrete choice probability, the pattern of region's limit distribution is closely related to the property of the externality effects of residents and firms. Zipf's pattern will emerge as the limit distribution if the externalities from resident and firm converge to an asymptotic value as city size getting large.

Hsin-Ping Chen  
Department of Economics  
National Chengchi University  
2001

**Keywords:** Zipf's Law, Gibrat's Law, Rank-Size Rule

## 1. Introduction

Zipf's law presents a robust regularity in cities' limiting distribution. The empirical evidence shows that countries with different economics structures mostly converge to the same limiting pattern. It is examined in most modern countries by Rosen and Resnick (1980); in India in 1911 by Zipf (1949); in U.S. history by Dobkins and Ioannides (1998), Krugman (1996) and Zipf (1949); and in mid-nineteenth century China by Rozman (1990). These empirical evidences of different countries and dates support Zipf's law.

This general distribution pattern can be viewed as a "self-organized criticality". Bak, Tang and Wiesenfeld (1987) had shown that dynamical systems might evolve into a "self-organized critical point". Zipf's law can be expressed as the probability that the size of a city is greater than some level  $S$  (or the rank of the city with city size  $S$ ) is proportional to  $1/S$ .

Gabaix (1999) states that Zipf's law is the necessary condition for a local growth model. Traditional explanations of Zipf's law mostly come from the economic force point of view (Losch 1954; Hoover 1954; and Beckman 1958). However, the balancing of transport cost, externalities and productivity difference do not really explain why the limit distribution is as Zipf's pattern (Gabaix 1999). Analysis based on random process prospect gives a very different explanation for this problem (Simon 1955). The assumption in Simon's model empirically shows

contradicting to the real world (Duncan Black and Vernon Henderson 1999). Gabaix (1999) proposes that the robust empirical regularity may due to the simple statistical property. He shows that a homogeneous growth process referred to as Gibrat's law would derive a limit distribution as Zipf's pattern. This finding connects the static limiting distribution to the corresponding dynamic growth process. Consequently, questioning the reason of Zipf's law could be related to examining the property of the growth process.

The purpose of this paper is to investigate the possible relation between the economic mechanism and the statistical condition (Gibrat's law) in explaining Zipf's distribution. We are interested in the following questions: Does a general economic model evolve to a city growth process characterized by Gibrat's law theoretically? Given the robust evidence of Zipf's law, what kind of implication that Gibrat's law could indicate in the economic growth model?

Section 2 introduces the Gibrat's law. Section 3 presents the possible relation between the growth process and Gibrat's law in a general economic model with and without a logit choice probability assumption. The conclusion is in Section 4.

## 2. Gibrat's Law

Gibrat's law was first presented by Robert Gibrat (1931) to model the relation

between the dynamics of firm size and industry structure. Sutton (1997) suggests that the evolution of market structure is a complex phenomenon and the proper understanding of the evolution of the structure may require analyzing both the economic mechanisms and the statistical effects.

The evolution of city growth is a complex phenomenon characterized by the features of complex systems. Gabaix (1999) proposed Gibrat's law to explain Zipf's law from a pure statistical view. Gibrat's law refers that cities with a homogeneous growth process would converge to Zipf's distribution in the steady state. A homogeneous random growth process has common mean and common variance which are independent of city size. He shows that an independent and identically distributed random growth process will have limit distribution with the tail distribution of city sizes at time  $t$  proportional to the inverse of city size in the steady state:

$$P(S > s) = \alpha / s \quad (1)$$

Where parameter  $S$  is the size of the city and parameter  $\alpha$  is a constant. This is Zipf's Law. Given a random growth process with the same mean and variance of growth rate, it could be proved that this Zipf's pattern is the necessary steady-state distribution.

Some empirical works have studied the property of the growth process of

cities. Eaton and Eckstein (1997) have found that there is no correlation between the initial size and the growth rate in both Japanese and French cities. Gabaix (1999) uses Eaton and Eckstein's data to show that the variance of the growth rate does not seem to differ across city sizes. We present the data in Taiwan to examine the relation between mean and variance growth rate versus city size in 1971 and 1998 in figure 1 and 2. There are 43 cities in the figures. The cities are chosen by the criterion that these are at least 50,000 inhabitants in the cities. The figures show that mean growth rate and the variance growth rate of cities in Taiwan between 1971 and 1998 are not strongly correlated with city size.

This Gibrat's law is a pure statistical explanation for Zipf's distribution. It would be important and legitimate to analyze the possible underlying economic mechanism and its correlation to this summarized statistical effect.

### 3. The Model

#### 3.1 A General Economic Model

Assume the population growth in a region is from migration and natural birth and death. There are overlapping generation residents with death rate,  $\kappa$ . Residents migrate to the city they choose from other region and live there until they die. Assume  $\alpha_{it}$  is the amenities level of city  $i$  at time  $t$ , and  $w_{it}$  is the wage in city  $i$  at time  $t$ . The amenities level is independent and identically distributed. The utility



level of a resident with consumption  $C$  in city  $i$  is  $Ca_{it}$ . Let  $S_{it}$  be the population in city  $i$  at time  $t$ . Assume there is external benefit (e.g., better public infrastructure and frequent social contact) and cost (e.g., congestion and air pollution) to the residents due to the size of the city:  $E(S_{it})$ . The migrants maximize the utility,  $E(S_{it})w_{it}a_{it}$ , to choose the residential city.

In equilibrium, the common utility level is:

$$U_t(w_{it}) = E(S_{it})w_{it}a_{it}, \quad (2)$$

Let  $M_t$  be the total immigrants in the region at time  $t$ . Let  $p_{it}$  be the probability that migrants will move to city  $i$  at time  $t$ . The migrant which move to city  $i$  at time  $t$ ,  $m_{it}$ , equals  $p_{it}M_t$ . The total population in city  $i$  at time  $t$ ,  $S_{it}$ , is:

$$S_{it} = p_{it}M_t + (1 + \tau - \kappa)S_{i,t-1}, \quad (3)$$

where parameter  $\tau$  is the birth rate in the region, and  $\kappa$  is the death rate in the region. Assume the production technology  $F$  is constant return to scale and the labor force is proportional to the population in the city. The production function in city  $i$  at time  $t$  is:

$$A(S_{it})F(S_{it}) = A(S_{it})F(m_{it} + (1 - \kappa)S_{i,t-1}) = A(S_{it})F(m_{it}, S_{i,t-1}) = A(S_{it})S_{i,t-1}f(m_{it}/S_{i,t-1}),$$

where function  $f$  is increasing and concave, and function  $A(S_{it})$  is the agglomeration effect of firms. This agglomeration effect is the external benefit and cost in firm's production (e.g., larger labor accessibility and frequent information

contact as benefit, and higher land rent and congestion as cost). The wages of the new immigrant is:

$$w_{it} = A(S_{it})f'(m_{it}/S_{i,t-1}) \quad (4)$$

The equilibrium utility level is:

$$U_t = a_{it}E(S_{it})A(S_{it})f'(m_{it}/S_{i,t-1}) \quad (5)$$

The growth rate of city  $i$  is as follows:

$$\gamma_{it} = (m_{it} + (\tau - \kappa)S_{i,t-1})/S_{i,t-1} = f'^{-1}(U_t/(a_{it}E(S_{it})A(S_{it}))) + (\tau - \kappa) \quad (6)$$

The equilibrium utility level,  $U_t$ , is the same for all cities in the region; and the amenities distribution is independent of the city size. If there is no external effect for residents and agglomeration effect in production, the growth rate of city is independent of city size. This growth process is referring to Gibrat's law. Consequently, the limiting distribution will converge to Zipf's pattern.

However, the existence of both external effect in resident,  $E(S_{it})$ , and agglomeration effect in production,  $A(S_{it})$ , suggests that the condition for the emergence of Zipf's pattern in the steady state is that the combination effect of these two external and agglomeration effects become independent of city size. However, these two external effects are mostly caused by city size.

The possible situation that this total external effect would be independent of city size is when these external effects reach the same boundary or criticality for

cities. Larger cities have external effects close to some boundary or criticality. This is corresponding to the empirical evidence that Zipf's law explains the limit distribution better for larger cities. If the net external effect (from positive and negative external effects) is bounded or have some common asymptotic value, the growth rate will become independent of size in the upper level. That is, Zipf's pattern will emerge in the limit distribution if the externalities from resident and firm converge to an asymptotic value as city size getting large.

Curry (1964) has mentioned that the resulting actions from different optimal decisions may appear as random as a whole. In another words, a deterministic rule may lead to a probabilistic result. This might due to the nonlinearly decision rule of agents, which inherent the property that deterministic rule may derive stochastic result.

### 3.2 A General Economic Model with the Logit Choice Probability

Assume the unsystematic part of the utility function is independent and identical extreme value. For a single random variable case, the extreme value distribution and the normal distribution are practically the same.<sup>1</sup> Given the extreme value distribution assumption, the location choice probability  $p_{it}$  is as logit model. The functional form and the discrete dynamic process of the choice probability are

---

<sup>1</sup> See Ch. 3, Kenneth Train (1990).

the following:

$$p_{it} = \frac{\exp(U_{it})}{\sum_j \exp(U_{jt})} \quad (7)$$

Take the rate of change of  $p_{it}$  with respect to time to derive the evolution process (Nijkamp and Reggiani, 1991).

$$\dot{p}_i = \dot{u}_i p_i (1 - p_i) - p_i \sum_{j \neq i} \dot{u}_j p_j \quad (8)$$

And then approximate this continuous evolution (equation (8)) in discrete time to derive the following discrete terms: (Wilson and Bennett, 1985)

$$p_{i,t+1} = (u_{it} + 1) p_{it} - u_{it} p_{it}^2 - p_{it} \sum_{j \neq i} u_{jt} p_{jt} \quad (9)$$

Where the variable,  $u_{it}$ , is the change of the utility level of residents in city  $i$  at time  $t$ .

The choice probability is influenced by it's own location benefit and the competitive forces from all other cities in the region. The gravity model has a logit type functional form of the location choice probability.

The growth rate of city size in city  $i$  at time  $t$  is:

$$\begin{aligned} \gamma_{it} &= (S_{i,t} - S_{i,t-1}) / S_{i,t-1} \\ &= \{p_{it} / [p_{i,t-1} + (1 + \tau - k) p_{i,t-2} + \dots + (1 + \tau - k)^{t-1} p_{i,0}]\} + (\tau - k), \end{aligned} \quad (10)$$

It shows that the growth rate of city  $i$ ,  $\gamma_{it}$ , is proportional to the corresponding location choice probability,  $p_{it}$ . Consequently, the property of the distribution of the city size growth rate is the same as the distribution of the growth rate of the choice probability.

Apply the discrete probability in equation (9) to derive the growth rate of the choice probability:

$$\theta_{it+1} = (p_{it+1} - p_{it}) / p_{it} = u_{it} - \sum_{j=1}^n u_{jt} p_{jt} \quad (11)$$

The average and variance of the growth rates of the choice probability in the steady state are the following:

$$E[\theta_{it+1}] = \mu_t - \sum_{j=1}^n \mu_{jt} E[p_{jt}] \quad (12)$$

$$Var[\theta_{it+1}] = \sum_{j=1}^n \mu_{jt}^2 Var[p_{jt}] + 2 \sum_{j < k} \mu_{jt} \mu_{kt} Cov(p_{jt}, p_{kt}) \quad (13)$$

The average growth rate of the choice probability is varied by the corresponding change of utility,  $u_{it}$ . In equilibrium, the utility level is the same across cities as in equation (5), the change of utility is the same across cities in the steady state. The average growth rate of the choice probability (equation (12)) and the growth rate of city size (equation (10)) are the same across cities.

Furthermore, the variances of the probability growth rate are also the same across cities. It shows that both the average and variances growth rates are independent of city size in the steady state. According to Gibrat's law, the general economic model with logit choice probability has limit distribution converge to Zipf's pattern.

#### 4. Conclusion

Gibrat's law explains Zipf's law in a very different way from traditional economics mechanisms. It is easy and neat as a pure statistical explanation. Using a general economic growth model with the logit choice probability assumption, the limit distribution will converge to Zipf's pattern without further conditions of the externalities according to Gibrat's law. Without the assumption of extreme value distribution or the logit choice probability, the presented general economic growth model shows that the properties of residential externalities and agglomeration effect closely affect the pattern of its limit distribution. The robust empirical evidence of Zipf's distribution implies that the externalities and agglomeration economies (from both resident and firm) are independent of city size in the steady state. Zipf's pattern will emerge in the limit distribution if the externalities from resident and firm converge to an asymptotic value as city size getting large.

Gibrat's law shows that a homogeneous growth process converges to Zipf's pattern in the steady state. The optimal migrant flows from gravity theory will generate a homogeneous growth process that is independent of city size. This growth process derived from agents' optimal solutions based on deterministic rule will have Zipf's limit distribution.

Two major sources of explanations for Zipf's law (economics mechanism and statistical method) are correlated vertically. The optimal decisions driven by the

economic forces are the microscope of agents' behavior, and the summarized effect of the growth process to the limit distribution as a whole in the region is at the macro level. Multiple optimal decisions in the micro level do not necessarily result in a deterministic result. On the contrary, it may appear random as a whole in the limit. A general location problem may endow the property of a complex system: deterministic rule leading to a stochastic result.

## References

- Bak, P., C. Tang and K. W., "Self-Organized Criticality: An Explanation of  $1/f$  noise", *Physical Review Letters*, Vol. 59, Number 4, 1987, 381-385.
- Beckman, M. "City Hierarchies and Distribution of City Size", *Economic Development and Cultural Change*, VI, 243-248, 1958.
- Black, Duncan and Henderson, Vernon. "A Model of Urban Growth", *Journal of Political Economy*, April 1999, 107(2), 252-284.
- Curry, L., "The Random Spatial Economy: An Exploration in Settlement Theory," *Annals of the Association of American Geographers*, LIV, 1946, 138-146.
- Dobkins, L., and Y. Ioannides, "Dynamic Evolution of the U.S. City Size Distribution," in J.M. Huriot and J.F. Thisse eds., *The Economics of Cities*, New York, NY: Cambridge University Press, 1998.
- Eaton, J., and Z. E., "City and Growth: Theory and Evidence from France and Japan", *Regional Science and Urban Economics*, XXVII (1997), 443-474
- Gabaix's, X., "Zipf's Law and the Growth of Cities," *American Economic Review Papers and Proceedings*, LXXXIX, 1999, 129-132.
- Gibrat, R., *Les Inegalites economiques* (Paris, France: Librairie du Recueil Sirey, 1931)
- Hoover, E.M., "The Concept of a System of Cities: A Comment on Rutledge Vining's Paper, *Economic and Cultural Change*, III, 196-198, 1954.
- Krugman, P., *The Self-Organizing Economy*, Cambridge, MA: Blackwell, 1996.
- Losch, A., *The Economics of Location*, Yale University Press, 1954.
- Marsili, M., and Y.C. Zhang, "Interacting Individuals Leading to Zip's Law", *Physical Review Letters*, Vol. 80, Number 12, 1998, 2741-2744.



Nijkamp, P. and A. Reggiani, Chaos Theory and Spatial Dynamics, *Journal of Transport Economics and Policy*, Vol. XXV, No. 1, 1991, 81-96

Rosen and Resnick, "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy," *Journal of Urban Economics*, VIII , 1980, 165-186.

Rozman, G., "East Asian Urbanization in the Nineteenth Century," in Van der Woude et al. , 1990, 61-73.

Simon, H. , "On a Class of Skew Distribution Function," *Biometrika*, XLII, 1955, 425-440.

Sutton, J., "Gibrat's Legacy", *Journal of Economic Literature*, XXXV (1997), 40-59.

Train, Kenneth, *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*, Cambridge, MA: MIT press, 1990.

Wilson, A.G. and R.J. Bennett, *Mathematical Methods in Human Geography and Planning*, J. Wiley & Sons, Chichester, New York, 1985

Zipf, G., *Human Behavior and Principle of Last Effort*, Cambridge, MA: Addison-Wesley, 1949.

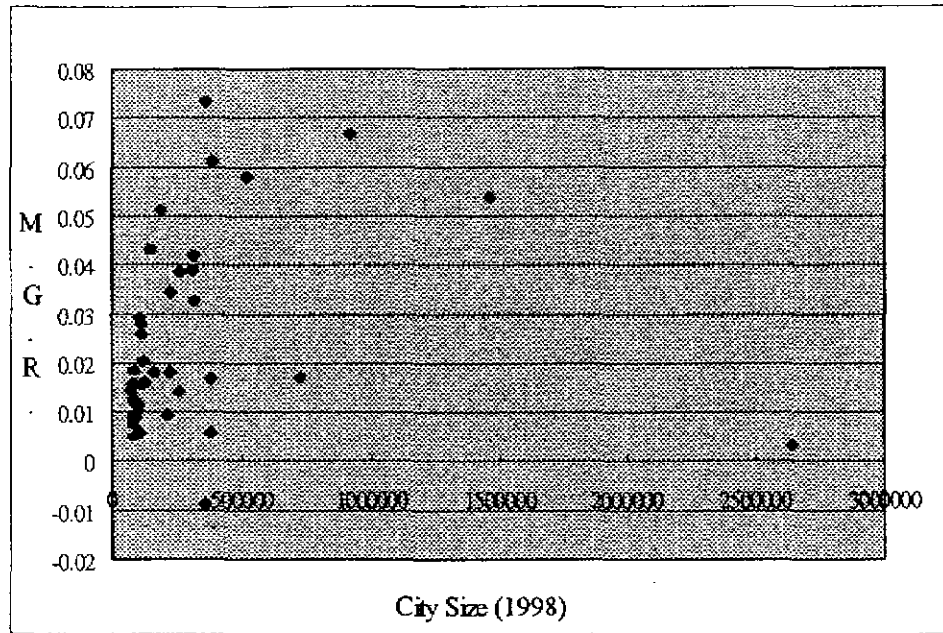


Fig. 1.1 The Mean Growth Rate (M.G.R.) versus City Size of Cities in Taiwan in 1998

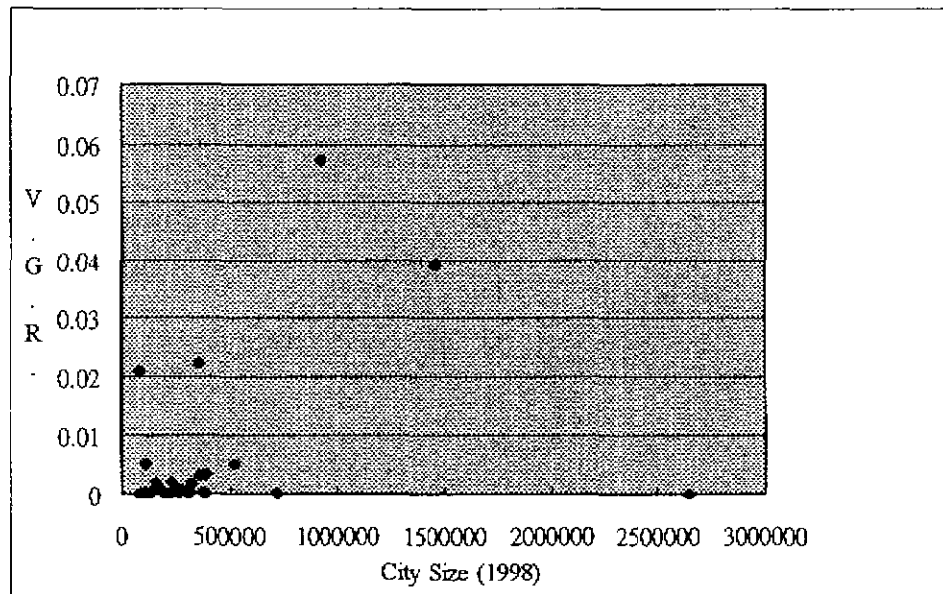


Fig. 1.2 Variance Growth Rate (V.G.R.) versus City Size of Cities in Taiwan in 1998

Source: Statistics Annals by Ministry of Interior

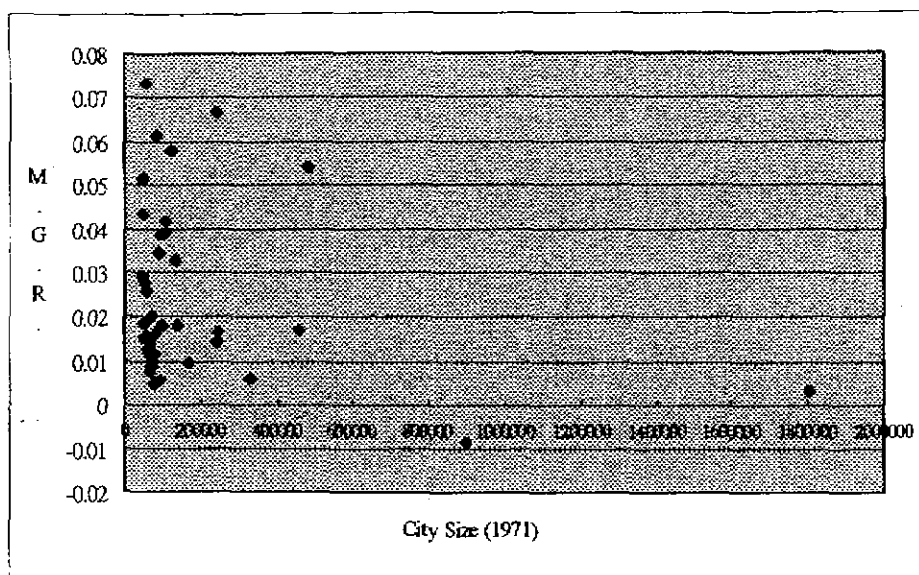


Fig. 2.1 The Mean Growth Rate (M.G.R.) versus City Size of Cities in Taiwan in 1971.

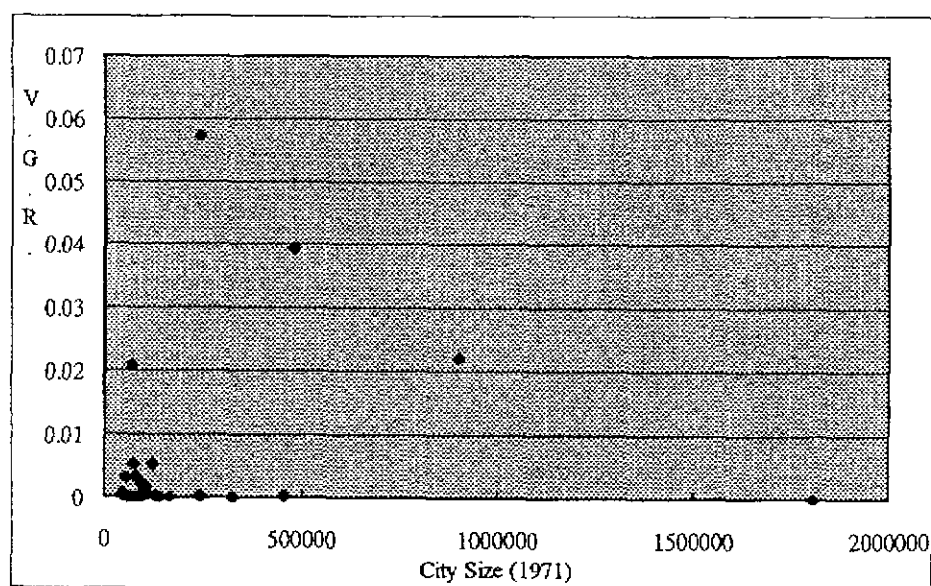


Fig. 2.2 The Variance Growth Rate (V.G.R.) versus City Size of Cities in Taiwan in 1971

Source: Statistics Annals by Ministry of Interior

# **The Theoretical Derivation and the Dynamic Process Simulation of the Rank Size Rule - Order from Random Growth Process in the Evolving Complex Systems**

## **Abstract**

Power law has been shown to be a common feature of many self-organized complex systems, and Zipf's law in regional science is the most famous of all these distributions. This paper shows that the assumption of homogeneity of the random growth process as assumed in Gibrat's law will generate city size distribution as power law. However, Gibrat's law does not necessarily generate Zipf's limiting pattern. City distribution could possibly converge to a Zipf's pattern limiting distribution only with a diminishing decreasing standard deviation of the random growth rate. Moreover, the value of the diminishing rate of the standard deviation of city growth rate determines the speed of the convergence and the value of the converged slope. The homogeneous random evolving process is the essential underlying feature, which generates the common power law property of many complex systems. Nevertheless, the variation of the changing rate of increased potential connections and the sensitivity of interactions among cities are the major reasons for the differences of the slopes among self-organized systems.

Hsin-Ping Chen  
Department of Economics  
National Chengchi University  
2001

**Keywords:** Self-organized criticality, Complex systems, Potential connections,

## 1. Introduction

The world is full of complex systems that are self-organized not only in response to exogenous disturbance, but also in response to internal logic. There is no global controller in the complex system. Many levels of hierarchical organization define the interaction among different level of units. Various deterministic interacted mechanisms in different level of hierarchy and the stochastic factors caused within and outside the system organize the systems' behavior. Global weather, developing embryo, global economy and growing cities are all examples of this self-organizing system. Global economy is composed of many levels of hierarchical organization that consists of various levels of agents and units interacted with each other. The evolution of the global economy results from the evolution of the system. Similarly, the growth of cities within a region, a country or globally is also corresponding to the self-organized mechanism. Cities' growth within a region is caused by both deterministic and stochastic factors. The deterministic factor is based on the systematic features of various hierarchical organizations. Those include the location choices of different agents (residents and firms) and the policy choices of different units; those choices are generated from systematic self-interest optimization decision process. Objects from various self-organizing systems, such as war, sandpiles, earthquakes, forest fires, and city distribution, display a size distribution as power function. Power law has been shown to be a common feature of many complex systems; this common feature

reflects certain regularity of the size distribution in different systems. Zipf's law in city distribution is the most famous regularity among all these distributions; it states a linear relation between log of city rank and log of city size of cities' limiting distribution in different countries. This feature indicates a certain limiting hierarchical composition of city distribution and is strongly supported by empirical evidence in different data set. The application of Gibrat's law explains Zipf's pattern by a statistical mechanism. However, this pure statistical explanation is lack of condition for the existence of the limiting distribution and also lack of consideration of the self-organized property and economic foundation.

The purpose of this paper is to investigate the possible underlying mechanisms or properties in complex systems that would lead to this general simple regularity given different structures. Furthermore, to explain power law and Zipf's law from the self-organized complex system point of view that has not been taking into account to explain Zipf's pattern before.

Section 2 introduces the essential idea and property of the dynamical complex systems. Section 3 gives a brief review of power law, rank-size rule and Zipf's law which describe urban size distribution in regional studies. Section 4 shows the common features of the self-organized criticality: power law. Section 5 investigates the emergence of power law and Zipf's law theoretically and empirically both from

the stochastic process and self-organized feature point of views. The conclusion is in Section 6.

## 2. The Evolving Complex System

Philip Anderson, the Nobel laureate physicist defined the complex system as a science of "emergence". It is about the surprising ensembles from the nonlinear combination of the interacting units in the system. From the structure point of view, the complex system is self-organized as systems form from almost random or homogeneous state based on some deterministic rules. Complexity is also defined as a measure of the sensitivity of particles in the systems. It is the potential connections among agents in the systems. The complex system is first observed in physics and biology. Further research about economy as an evolving complex system is discussed in Santa Fe Institute in 1987. The nonlinear dynamic structure among the interacting units results in several special features.

Author, Durlauf and Lane (1997) pointed out six features of the complex system: no global controller, dispersed interaction, crosscutting hierarchical organization, continual adaptation behaviors, perpetual novelty niches, and out-of-equilibrium dynamics. The organized units in the systems are interacted with each other in a very pluralistic form. There are many levels of hierarchical organization

defining the interaction among different level of units. Various deterministic interaction mechanisms in different level of hierarchy organize the systems' behavior. There is no global controller in the system. In addition, the various deterministic interaction rules are adaptive, and the system is perpetual novelty. The most important feature that is very much different from the classic economic theory is the out of equilibrium features. There is no unique optimum or global equilibrium in the system.

Systems with these features are called adaptive nonlinear networks. Due to the structural features of the system, there are some essential properties of this adaptive nonlinear network. The possible evolving outcomes of the system is path dependent; the historical shock is very crucial for the result. Also, it shows "lock-in" property in the evolution process. Once an "alternative" is chosen, it is difficult to exit (named as lock-in effect) due to the increasing return to scale property. Thus, the evolving result is possible to be inefficient as the whole or at the end. Even there is a better alternative later, the system might stock in the less efficient alternative due to the increasing return to scale property. The outcome of the system is not predictable and multiple equilibria are possible.

The evolving outcome is very sensible to the initial conditions and the value of parameters. Different value of the parameters will result in different characteristic of the evolving process. Some evolving process is predictable and stable, but some are



unpredictable and unstable. Features of this self-organized complex system have been applied to study various economic phenomena, such as positive feedback in economics, the historical path dependence in urban systems, nonlinear theory in global business cycle, input and output structure in the percolation economics, and the financial feedback in market mechanisms ( see Arthur 2000; Krugman 1996; Day 1994). Urban systems are composed of different levels of hierarchy units whose location and policy decisions form the systematic part of the system behaviors.

Krugman (1996) suggested two principles of self-organized process in explaining economic system: order from instability and order from random growth. The first principle indicates the astonishing empirical evidence that simple order emerges from the unstable self-organized criticality that is generated from the out-of-equilibrium dynamical systems. The second principle states the possible explanation for the feature of the emergence of order. Gibrat's law is a typical example. This principle will be further investigated.

### 3. Zipf's law and Gibrat's law

Pareto distribution is most commonly employed to study urban sizes in regional research (see Mills and Hamilton 1994):

$$G(x) = Ax^{-\alpha} \quad (1)$$

where  $G(x)$  is the number of cities with at least  $x$  people; it could be interpreted as the rank of the city with  $x$  people. Variable  $x$  is the city size; parameters  $A$  and  $\alpha$  are constants to be estimated from the data (cited from Mills and Hamilton 1994, p. 78). Relation in equation (1) is also called power law, which describes the number of cities with a population larger than  $x$  is proportional to  $x^{-\alpha}$  (see Fujita, Krugman and Venables 1999, ch. 12). The power law with the exponent,  $\alpha$ , close to one, is referred to as Zipf's Law. The alternative name is the rank-size rule (see Fujita, Krugman and Venables 1999, p.217):

$$G(x) = Ax^{-1} \quad (2)$$

It refers the following relation:

$$xG(x) = A \quad (3)$$

Zipf's law (or rank-size rule) proposes that the product of the city size ( $x$ ) and its rank ( $G(x)$ ) is a constant. Take the log of the city size and city rank in equation (2):

$$\ln(Rank) = A' - \ln(Size) \quad (4)$$

where  $A' = \ln(A)$ ; it is a constant. Equation (4) is an alternative expression and most commonly employed to present Zipf's law (or rank-size rule). This relation has robust empirical evidence across countries and time. When an urban system has city distribution as Zipf's law, the constant  $A$  in equation (3) represents the population of the largest city in the region. Consequently, the second-largest city would have one-

half the population of the largest, and the third, one-third that population, etc. This distribution regularity in various urban systems are based on different levels of economic decision processes.

Gabaix (1999) proposes Gibrat's law to describe Zipf's pattern in city distribution. He states that if cities grow randomly with the same expected growth rate and the same standard deviation, the limiting distribution will converge to Zipf's law (Gabaix 1999). Gibrat's law indicates that when cities' size randomly grows with a homogeneous growth process<sup>1</sup>, the limiting distribution will converge to Zipf's pattern regardless of the distribution type of the growth rate.

#### 4. The common order from the self-organized criticality: power law

The hierarchical structures and the interaction among agents generate the complex evolution process. Per Bak et al. demonstrated that dynamical systems naturally evolve into highly interactive critical states which are barely stable ( see Bak, Tang, and Wiesenfeld 1987 ). This self-organized criticality, where a small cause may lead to a large event, is the common underlying mechanism behind the special feature: the power law distribution of the corresponding events ( see Bak and Paczuski 1995 ). The events, named as the complexity cascades of self-organized criticality in social

---

<sup>1</sup> The expected growth rate and standard deviation of city growth rate are the same across cities; they are independent of city size

system, take the form of wars, strikes, economic depressions, collapses of government, coalitions, emergence of cities, and many others. The historical details of these events are unpredictable; however, the statistical distribution of these events is predictable. This predictable power law appears as a straight line in a double logarithmic plot of rank and size. It is empirically observed in various complex systems: ecological systems, social systems and geophysical phenomena. They are organized by various mechanics and networks type, but all results in characteristic power function in the size distribution.

Brunk (2000) propose that the evolution process of the complex system characterized by the self-organized criticality is composed of two parts: a systematic growth factor and a random growth factor. The systematic growth factor is characterized by the underlying mechanism and depends on the current degree of "complexity" of the systems. The complexity is measured as the degree of the potential connection among members, and is growing as the systems getting larger or extending more subsystems. The increasing complexity implies increasing in the sensitivity of interactions. A random growth factor could be generated either inside or outside the systems. These systematic and unsystematic growth factors together with the increasing complexity property in the complex systems are essential in explaining power function phenomenon. The most robust evidence of power function in

complexity cascades appears in urban system.

Zipf (1949) found that city sizes follow a very simple distribution law. He showed that the size distribution of cities satisfies power law with a scaling exponent equals one empirically for many different societies and time periods. It is named as Zipf's law showing that the probability the size of a city greater than some size  $x$  is proportional to  $1/x$ . The expression of Zipf's law can be visualized by taking a country cross sectional data on city size and city rank. The plot with the log of rank along the y-axis, and the log of the population along the x-axis will mostly show a straight line with slope very close to  $-1$ .

The empirical evidence of Zipf's law is shown in various data sets: most modern countries ( see Rosen and Resnick 1980 ), India in 1911 ( see Zipf 1949 ), U.S. history ( see Dobkins and Ioannides 1998; Krugman 1996; Zipf 1949 ) and China in the mid-nineteenth century ( see Rozman 1990 ). More recent data shows that Zipf's law remains rather good approximation for developed countries. However, cities in countries with a unique social structure, such as the former USSR or China, do not quite follow Zipf's law ( see Marsili, and Zhang 1998 ).

Empirical evidences show that individual decision rules followed by the underlying hierarchical structure do influence the generating process of self-organized criticality. The growth of the cities in the region is characterized by the hierarchical

system and is self-organized. Gabaix (1999) proposes Gibrat's law to describe Zipf's law in city distribution. However, Gabaix's work only shows that if there is a steady state process of the tail distribution of city sizes, the limiting distribution would be as Zipf's pattern. He did not discuss the condition for the existence of the steady state process.

This finding of Gibrat's law suggests a pure statistical view in explaining Zipf's law other than the usual economics views based on the urban theories. According to the urban theories, agents in the space system (residents and firms) make their location decisions depending on their corresponding economics considerations to optimize their own goals. The major consideration in location decisions includes the location accessibility and the agglomeration effects. The agglomeration effect could be measured as the sum of agents (residents or firms) in each city weighted by its corresponding location accessibility. The larger the agglomeration effects the larger the location attractiveness; consequently, the faster this place would grow. According to this explanation, it seems counter intuitive that city growth is independent of city size as Gibrat's law proposing. In the next section, we will investigate the possible explanation for the Zipf's law concerning self-organization features, location theory and Gibrat's law.

## 5. A model and the simulation about power law

### 5.1 The model

According to Gibrat's law (Gabaix 1999), Cities' limiting distribution will converge to Zipf's pattern if they grow randomly with the same expected growth rate and the same standard deviation.

One way to present the distribution of randomly growing cities is by Joseph Steindl (1965): Consider a region consists of  $n$  zones. Let  $x_{i,t}$  denotes the normalized size of zone  $i$  at time  $t$ , which is the population of city  $i$  divided by the total region population. The growth of city size is proportional to the current city size. The growth rate of zone  $i$  at time  $t$ ,  $\varepsilon_{i,t}$ , is identically independent distributed with mean *zero* and variance  $\sigma_{i,t}^2$ . The growth process of the normalized size of city  $i$  at time  $t$  is as follows:

$$x_{i,t} - x_{i,t-1} = \varepsilon_{i,t} x_{i,t-1}. \quad (5)$$

$$x_{i,t} = (1 + \varepsilon_{i,t}) x_{i,t-1} = x_{i,0} (1 + \varepsilon_{i,1}) (1 + \varepsilon_{i,2}) \dots (1 + \varepsilon_{i,t}) \quad (6)$$

The sum of the normalized size of all cities in the region at certain time period is one:

$\sum_i x_{i,t} = 1$ . Consequently, the average normalized size at each time period stays

constant through time. Assuming the time period is so short that the growth rates,  $\varepsilon_{i,t}$ , is relatively small. This justify the following approximation:

$$\log(1 + \varepsilon_{i,t}) \cong \varepsilon_{i,t}. \quad (7)$$

Taking logs of equation (6); it becomes:

$$\log x_{i,t} = \log x_{i,0} + \varepsilon_{i,1} + \varepsilon_{i,2} + \dots + \varepsilon_{i,t} \quad (8)$$

The term  $\log x_{i,0}$  would be very small comparing to the term  $\log x_{i,t}$  as  $t \rightarrow \infty$ , the distribution of  $\log x_{i,t}$  is approximately normal distribution with mean *zero* and variance  $\sigma_t^2 t$  when the growth rate of cities at time  $t$  have the same variance  $\sigma_t^2$  across cities. Thus, the distribution of normalized city size  $x_{i,t}$  is lognormal.

In addition, city rank can be expressed as the tail distribution of city sizes at time  $t$ :

$$R_t(X) = P(x_t > X) \quad (9)$$

As  $t \rightarrow \infty$ , the observations of  $\log R(X)$  versus  $\log X$  would lie in a straight line with negative slope if the distribution of normalized city size were lognormal. This represents the power pattern limiting distribution:

$$R(X) = \alpha / X^\beta \quad (10)$$

This is the power law feature that is common in many evolving complex systems. Generally, a group of entities with homogeneous random growth rate will generate power law limiting distribution if there is steady state process.

An alternative method in Gabaix (1999) is applied here to investigate the possible limiting distribution given a homogeneous random growth process. Assume the city growth process is as in equation (5). It follows:

$$x_{i,t} = (1 + \varepsilon_{i,t})x_{i,t-1} = \gamma_{i,t}x_{i,t-1} \quad (11)$$

where  $\gamma_t = 1 + \varepsilon_t$ , and  $x_{i,t}$  denotes the normalized size of zone  $i$  at time  $t$ . The tail



distribution of city sizes at time  $t$  (equation (9)) can be expressed as (Gabaix, 1999):

$$R_t(X) = P(x_t > X) = \int_0^\infty R_{t-1}\left(\frac{X}{\gamma_t}\right) f(\gamma_t) d\gamma \quad (12)$$

Assume a general distribution of the type  $R_t(X) = \alpha/X^{\beta(t)}$ , we could derive a general term from equation (12):

$$R_t(X) = \alpha X^{-\beta(0)} E[\gamma_1^{\beta(0)}] E[\gamma_2^{\beta(1)}] \dots E[\gamma_t^{\beta(t-1)}], \quad (13)$$

The condition of the existence of a steady state process  $R_t(X) = R(X)$  is:

$\lim_{t \rightarrow \infty} E[\gamma_t^{\beta(t-1)}] = 1$ . A constant,  $\beta(t) = \beta$ , will do. The value of parameter  $\beta$  is not

necessarily to be one to assure the existence of a steady state limiting distribution.

This indicates that if there is steady state process, the limiting size distribution is as power function, not necessarily as Zipf's law. This finding is different from statement of Gabaix (1999) that the limiting distribution would be Zipf's law.

## 5.2 Simulation

The simulation experiments in this section is based on the homogeneous growth process assumed in Gibrat's law. Cities randomly grow with the growth rate of the same expected mean and standard deviation. The growing process is independent of city size. A homogeneous growth process across cities includes two types: (1) homogeneous across cities and time, and (2) homogeneous across cities only. Both types of growth process will be simulated in this section.

In experiment 1, we assume the first type of growth process that cities grow

with the same expected mean and standard deviation across cities and time. The standard deviation of the random growth factor is assumed constant across cities and time:  $\sigma_{i,t}^2 = \sigma^2$ , where  $\sigma_{i,t}^2$  represents the variance of the growth rates of city  $i$  at time  $t$ . The variance of the random growth rate indicates the degree of growth stability of certain city at certain time. If cities in the region grow with different variance ( $\sigma_{i,t}^2$ ), it shows that cities grow with different degree of stability. Cities with smaller variance of growth rate will have growth pattern more stable than cities with larger variance of growth rate. If cities in the urban system grow with the same variance ( $\sigma_i^2$ ), the variance of the random growth rate represents the degree of the heterogeneity among cities' growth. The larger the variance, the more diverse cities grow; on the contrary, the smaller the variance, the more evenly cities grow.

The evolutions of cities in the urban system are interacted with both competitive and cooperative relationships. The growth of city in the urban system reflects the location advantage from the interaction among cities in the system. Thus, the heterogeneity among cities' growth mainly refers to the relative location advantages or disadvantages among cities. A constant variance of the random growth rate across cities and time ( $\sigma^2$ ) indicates that the differences among cities' location attractiveness remain the same across time. However, in the real world, the differences among cities' location attractiveness tend to change across time due to the

change of various economic phenomenon. We simulate this urban system and estimate the slope of equation (14) to study its evolving process.

$$\ln(Rank) = A' + \alpha \ln(Size) \quad (14)$$

This relation is derived from pareto distribution in equation (1) of cities distribution. The sign of the slope ( $\alpha$ ) is negative due to power law; the absolute value of the slope indicates the degree of diverse in cities' size. A larger absolute value of the slope implies that size of cities is more evenly distributed; a smaller absolute value of the slope implies a more heterogeneous size distribution.

In this experiment, we simulate city growth given homogeneous random growth rate across cities and time. The parameter values and the simulation processes are presented in Appendix. The estimated slopes ( $\alpha$ ) in equation (14) given different value of standard deviation of the random growth rate are presented in Table 1. All the estimated slopes are negative as pareto distribution describes. All estimated absolute value of slopes given different value of standard deviation decrease as time pass. This denotes that cities in the urban system evolve into more and more heterogeneously distributed. In other words, sizes of cities are getting more and more diverse as time pass by. Furthermore, the estimated absolute value of the slope is getting smaller as time pass without a limit boundary. This result indicates that cities would continue heterogeneously grow without a limiting distribution given the homogeneous random

growth rate across locations and time. A constant slope indicates a power law distribution. As the slope equals -1, the distribution fulfills Zipf's law.

This finding is inconsistent with Gibrat's law. Given the assumption of Gibrat's law, Gabaix only shows that if the steady state exists, Zipf's pattern would be the limiting distribution in the steady state. In other words, he shows that Zipf's limiting distribution is the possible result if there is steady state given the homogeneous growth process. Gibrat's law does not assure that Zipf's distribution is the necessary condition of the homogeneous growth process. The simulation shows that the assumption in Gibrat's law does not assure the existence of the steady state in size distribution; however Zipf's pattern emerges only when a steady state process exists. We would like to discuss further the possible condition for the existence of a steady state process given the assumption in Gibrat's law.

As we mentioned earlier, cities' random growth rate with a smaller standard deviation ( $\sigma^2$ ) denotes less difference among cities' location attractiveness in the region; thus, cities' growth is less various; cities' size is more evenly distributed. The less the differences of the location attractiveness among cities, the larger the absolute value of the regression slope ( $\alpha$ ). On the contrary, a random growth rate with a larger standard deviation indicates larger differences among cities' location attractiveness in the region; cities' growth is more diverse; cities' size is more heterogeneously

distributed; and the estimated slope is flatter (Table 1).

In this evolving urban system, agents choose cities to locate according to their decision rule one by one and period by period. There is no global controller designing the growth of the region. It is all determined by the dispersed interaction among agents. Also, according to the adaptation behaviors within the hierarchical structures in the system, the potential connections and interactions among agents and locations are getting more and more frequent and sensitive during the evolution process. More frequent interactions and increasing sensitivities among locations and agents would lead to less variance of the relative location attractiveness among cities. Thus, this growing sensitivity feature could be characterized by a decreasing standard deviation of cities' random growth rate. However, the increase of the sensitivities and connections among agents and locations are reducing and converging to a critical state. This feature could be characterized as a diminishing decrease of the standard deviation of cities' random growth rate.

Cities grow based on the growth rate with a diminishing decreasing standard deviation is simulated in experiment 2. The simulation results are presented in Table 2. Assume that the standard deviation of the random growth factor,  $\sigma_t$ , is decreasing in a reducing rate across time. We use a simple form to express this standard deviation:

$$\sigma_t = \sigma_0 t^{-b} \quad (15)$$

Where the parameter,  $b$ , represents the reducing rate of the standard deviation. This equation is a simple model reflecting a diminishing decreasing standard deviation. A growth rate with a constant standard deviation across cities and time will generate an urban system with cities evolving into more and more heterogeneously distributed without limiting distribution; the absolute value of the slope in equation (14) is decreasing. A growth rate with a decreasing standard deviation without boundary across time will generate an urban system with cities evolving into less and less heterogeneously distributed in a constant rate; the absolute value of the slope in equation (14) is decreasing with a limiting value. A growth rate with a diminishingly decreasing standard deviation across time will generate an urban system with cities evolving into less and less heterogeneously distributed in a diminishing rate; the absolute value of the slope in equation (14) is decreasing to a limiting value in a diminishing rate.

Given the same initial condition, a larger diminishing rate ( $b$ ) denotes the standard deviation,  $\sigma_t$ , decreases faster across time. The faster the standard deviation decreases, the faster the dispersion of the growth rate among cities is reduced; consequently the slower the speed of changing into heterogeneous city size distribution is; finally, the faster the urban system converge to a limiting distribution. By the way, the larger the diminishing rate of the standard deviation, the smaller the

value of the standard deviation, and the larger the absolute value of the slope. On the contrary, a smaller diminishing rate indicates that cities' size distribution is more heterogeneous distributed and converge slower. These features are observed in Table 2.

The random growth rate with a diminishing decreasing standard deviation across time affects the reduction speed of the regression slope. As the potential connection and sensitivity among cities increase across time, the diverse of the relative location advantage among cities decrease, and consequently the variance of the growth rate across cities decrease. The speed of cities' evolution from uniform to heterogeneous is reducing. As the potential connections and sensitivity of interactions among cities are increased in a diminishing rate, the speed of cities' evolution into heterogeneous distribution is reduced and hopefully the urban system converge to a limiting pattern given certain parameter value. Furthermore, the converged slope is closely related to the speed of the decreasing rate of the variance of the growth rate.

Simulation result in Table 2 shows that a diminishing decreasing standard deviation of the growth rate under Gibrat's law possibly lead to a convergence of size distribution given certain parameter value. That is, this further condition of the standard deviation of growth rate given Gibrat's law may result in a steady state

process. However, it does not promise the limiting pattern in the steady state process as Zipf's pattern. Zipf's pattern only emerges given certain values of parameters. We could conclude that given the assumption in Gibrat's law, a diminishing decreasing standard deviation of the growth rate is the necessary condition for the existence of a steady state process and a limiting distribution.

## 6. Conclusion

Power law has been shown to be a common feature of many complex systems, and Zipf's law in regional science is the most famous of all these distributions. The application of Gibrat's law explains Zipf's pattern by a statistical mechanism. However, this pure statistical explanation is lack of condition for the existence of the limiting distribution and considerations from self-organized process. This paper is the first work investigating the source of Zipf's law from the complex system point of view. It shows that if region grows based on random growth rates with the same mean and variance across cities, it will generate power law distribution of city size. However, a random growth process as assumed in Gibrat's law does not necessarily generate Zipf's limiting pattern. According to the features of the dynamic complex systems, the adaptation behaviors of the interacted agents under the hierarchical structures in the system will lead to more frequent potential connections among agents and higher sensitivity of interactions. As the potential connections and sensitivity of



interactions among agents and cities increase into a critical state across time, it reflects the difference of the location attractiveness among cities in the region is reducing in a diminishing rate. This could be characterized by a growth rate with diminishing decreasing standard deviation. The simulating result shows that a diminishing decreasing standard deviation of the random growth rate could possibly generate a Zipf's limiting distribution given certain parameter value. Moreover, the value of the diminishing rate determines the speed of the convergence and the value of the converging slope.

Generally, this finding also explains the power law features of the other complexity cascades from the corresponding self-organized complex systems. System particles with homogeneous random evolving rate would generate the power function distribution. A homogeneous random evolving process is the essential underlying feature, which generates the common power law property of many complex systems. Nevertheless, the major reason that the slopes of power functions differ among various self-organized critical systems is the variation of the changing rate of the increased potential connections and sensitivity of interaction within the systems. In other words, the converging slope of the power function is highly related to the changing rate of the increment of potential connection and the level of complexity.

Our finding suggests that homogeneous growth process does not assure the

existence of the Zipf's pattern distribution. Cities grow with the random growth process with diminishing decreasing standard deviation would lead to a convergence of city distribution. The diminishing rate and the initial value of the standard deviation would affect the pattern of the limiting distribution. That is, Zipf's limiting distribution will appear given a diminishing decreasing standard deviation growth rate with certain value of parameters.

## Reference

Author, W., *Increasing Returns and Path Dependence in the Economy*, Ann Arbor, the University of Michigan Press, 2000.

Author, W., Durlauf, S. and D. Lane, *The Economy as an Evolving Complex System II*, Reading, Mass.: Addison-Wesley, Advanced Book Program, 1997.

Bak, P. and M. Paczuski, "Complexity, contingency, and Criticality", *Proc. Natl. Acad. Sci.*, USA, Vol. 92, p.6689-8896, 1995.

Bak, P., Tang, C., and K. Wiesenfeld, "Self-Organized Criticality: An Explanation of  $1/f$  Noise", *Physical Review Letters*, Vol. 59, Num. 4, p.381-384, 1987.

Brunk, Gregory G. "Understanding Self-Organized Criticality as a Statistical Process." *Complexity*, Vol. 5, No. 3, p.26-33, 2000.

Day, R.H., *Complex Economic Dynamics, Volume I: An Introduction to Dynamical Systems and Market Mechanisms*, The MIT Press, Cambridge, Massachusetts, 1994

Dobkins, L., and Y. Ioannides, "Dynamic Evolution of the U.S. City Size Distribution," in J.M. Huriot and J.F. Thisse eds., *The Economics of Cities*, New York, NY: Cambridge University Press, 1998.

Fujita, M., Krugman, P., and A. J. Venables, *The Spatial Economy - Cities, Regions, and International Trade*, The MIT Press, Cambridge, Massachusetts, 1999.

Gabaix, X. " Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics*, Vol. 114 (3), p739-67, August 1999.

Krugman, P., *The Self-Organizing Economy*, Cambridge, MA: Blackwell, 1996.

Marsili, M. and Y.C. Zhang, "Interacting Individuals Leading to Zipf's Law," *Physical Review Letters*, Vol. 80, Num. 12, p.2741-2744, 1998.

Mills, E. and B. Hamilton, *Urban Economics*, Harper Collins College Publishers, Fifth edition, 1994.

Rosen, K., and M. Resnick, "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy," *Journal of Urban Economics*, VIII, 165-186, 1980.

Rozman, G., "East Asian Urbanization in the Nineteenth Century," in *Van der Woude et al.*, 1990, 61-73.

Steindl, J. *Random processes and the growth of firms: A study of the pareto law*. London: Griffin, 1965.

Zipf, G., *Human Behavior and Principle of Last Effort*, Cambridge, MA: Addison-Wesley, 1949.

Table 1  
Estimated slope ( $\alpha$ ) in Simulation of Cities Growth with Constant Variance ( $\sigma^2$ )

$$\ln(Rank) = A' + \alpha \ln(Size)$$

Experiment	(1)	(2)	(3)
$\sigma$	0.01	0.05	0.1
S1	-10.81	-2.39	-1.36
S2	-7.96	-1.70	-0.82
S3	-7.17	-1.33	-0.67
S4	-6.12	-1.12	-0.59
S5	-5.10	-1.16	-0.53
S6	-4.59	-1.02	-0.50
S7	-4.31	-0.99	-0.44
S8	-3.95	-0.87	-0.42
S9	-3.74	-0.84	-0.41
S10	-3.57	-0.79	-0.39

*Notes:*

The number of cities: 100.

The experimental time periods: 500.

$\sigma$  : The standard deviation of the random growth rate.

S1: The estimated slope of the regression of log rank against log size at time  $t=50$ .

S2: The estimated slope of the regression of log rank against log size at time  $t=100$ .

S3: The estimated slope of the regression of log rank against log size at time  $t=150$ .

S4: The estimated slope of the regression of log rank against log size at time  $t=200$ .

S5: The estimated slope of the regression of log rank against log size at time  $t=250$ .

S6: The estimated slope of the regression of log rank against log size at time  $t=300$ .

S7: The estimated slope of the regression of log rank against log size at time  $t=350$ .

S8: The estimated slope of the regression of log rank against log size at time  $t=400$ .

S9: The estimated slope of the regression of log rank against log size at time  $t=450$ .

S10: The estimated slope of the regression of log rank against log size at time  $t=500$ .

Table 2

Estimated slope ( $\alpha$ ) in Simulation with Decreasing Variance ( $\sigma_t^2$ )\*

$$\ln(Rank) = A' + \alpha \ln(Size)$$

Experiment	(1)	(2)	(3)
$\sigma_0$	0.1	0.1	0.1
$b$	1	0.23	0.1
S1	-5.918	-1.881	-1.543
S2	-5.903	-1.566	-1.181
S3	-5.898	-1.420	-1.116
S4	-5.877	-1.283	-0.989
S5	-5.860	-1.203	-0.884
S6	-5.856	-1.208	-0.805
S7	-5.841	-1.164	-0.723
S8	-5.839	-1.121	-0.664
S9	-5.843	-1.097	-0.611
S10	-5.836	-1.097	-0.593

$$*\sigma_t = \sigma_0 t^{-b}$$

Notes:

The number of cities:100.

The experimental time periods: 500.

S.D.: The standard deviation of the random growth rate.

S1: The estimated slope of the regression of log rank against log size at time t=50.

S2: The estimated slope of the regression of log rank against log size at time t=100.

S3: The estimated slope of the regression of log rank against log size at time t=150.

S4: The estimated slope of the regression of log rank against log size at time t=200.

S5: The estimated slope of the regression of log rank against log size at time t=250.

S6: The estimated slope of the regression of log rank against log size at time t=300.

S7: The estimated slope of the regression of log rank against log size at time t=350.

S8: The estimated slope of the regression of log rank against log size at time t=400.

S9: The estimated slope of the regression of log rank against log size at time t=450.

S10: The estimated slope of the regression of log rank against log size at time t=500.

## Appendix

### 1. Experiment 1 (Table 1): Cities' growth rate with constant standard deviation

#### *Simulation parameter:*

Number of cities in the region: 100

The experiment time periods: 500

The observed time periods: 50

The initial city size: 1

Cities in the region randomly grow with normal distribution

(Mean=0, Variance =  $\sigma^2$ ,  $\sigma$  =0.01, 0.05, and 0.1)

#### *Simulation process:*

1. Generate growth rate by randomly drawing from a normal distribution with mean zero and each of the assumed constant standard deviation (0.01, 0.05, and 0.1)
2. In period 1, 100 Cities grow according to the randomly drawing growth rate in step 1 given the same initial population (initial city size =1). Derive 100 cities size at period 2.
3. In period 2, 100 Cities continue growing from period 1 according to the randomly drawing growth rate in step 1. Derive 100 cities size at period 2.
4. Repeat step 3 until period 50; estimate the slope by regressing  $\ln(\text{rank})$  versus  $\ln(\text{size})$ .
5. Repeat step 4 until period 500.

2. Experiment 2 (Table 2): Cities' growth rate with diminishing decreasing standard deviation

*Simulation parameter:*

Number of cities in the region: 100

The experiment time periods: 500

The observed time periods: 50

The initial city size: 1

Cities in the region randomly grow with normal distribution

(Mean=0, Variance =  $\sigma_t^2$ ,  $\sigma_t = \sigma_0 t^{-\alpha}$ )

( $\sigma_0, \alpha$ )=( 0.1, 1 ), ( 0.1, 0.23 ) and (0.1, 0.1)

*Simulation process:*

1. Generate growth rate by randomly drawing from a normal distribution with mean zero and each value of the assumed decreasing standard deviation

$$(\sigma_t = \sigma_0 t^{-\alpha})$$

2. In period 1, 100 Cities grow according to the randomly drawing growth rate in step 1 given the same initial population (initial city size =1). Derive 100 cities size at period 2.

3. In period 2, 100 Cities continue growing from period 1 according to the randomly drawing growth rate in step 1. Derive 100 cities size at period 2.

4. Repeat step 3 until period 50; estimate the slope by regressing  $\ln(\text{rank})$  versus  $\ln(\text{size})$ .

5. Repeat step 4 until period 500.