Physica A 388 (2009) 2020–2024

Contents lists available at ScienceDirect

Physica A

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

Zipf distribution in top Chinese firms and an economic explanation

Jianhua Zhang, Qinghua Chen *, Yougui Wang

Department of Systems Science, School of Management, Beijing Normal University, Beijing 100875, People's Republic of China

ARTICLE INFO

Article history:
Received 3 September 2008
Received in revised form 15 January 2009
Available online 6 February 2009

PACS:
89.65.Gh
89.75.Da
05.90.+m

Keywords:
Firm size distribution
Zipf's law
AK model
Multiplicative process

ABSTRACT

By analyzing the data of top 500 Chinese firms from the year 2002 to 2007, we reveal that their revenues and ranks obey the Zipf's law with exponent of 1 for each year. This result confirms the universality of firm size character which has been presented in many other empirical works, since China possesses a unique ideological and political system. We offer an explanation of it based on a simple economic model which takes production and capital accumulation into account.

1. Introduction

Firms or enterprises, responsible for carrying out economic activities by providing various products and services, play key roles in countries. Firm size can be measured by a number of quantities: revenue, asset, employee or profit and so on. It is often used to represent firm's competition capability in the economy. Firm size varies enormously, from the sole trader up to a huge corporation employing thousands of people. For instance, *Wal-Mart Stores*, according to *Fortune* magazine of May 5th, 2008, ranked 1st in the top 500 American corporations list again for its amazing size. The value of its asset, revenue and profit is respectively 163,514 million, 378,799 million and 12,731 million dollars. The size distribution of firms in many countries was found to be likely consistent—it is highly skew, the small firms are extremely common, whereas large ones are extremely rare [1]. This topic has been attractive all along, especially in testifying its validity in various economies and seeking theoretical explanations.

As a pioneering study on the size distribution of firms, Gibrat reported that firm size could be described by lognormal distribution [2]. The same distribution was discovered in some followed empirical studies [3–6]. However, other empirical studies show that the size distribution (especially at upper tail) can be approximated closely by the Pareto or power-law distribution in the form of $P(X > x) \sim x^{-k}$ [7–9]. Inspired by the well-known work of Zipf [10], who firstly found that assets of US corporations follow Zipf's law: $S(r) \sim r^{-\beta}$, where $S(r)$ is the size of the corporation ranked r in a list ordered by asset size, beginning with the largest, great interest was attracted in the relationship of sizes and their corresponding ranks, and many considerable evidences on Zipf distribution of firm size were discovered later [11–14]. Actually, Zipf's law is consistent with Pareto or power-law distribution, they are different representations of the same fact. The relation of their exponents can be given by $\beta = 1/k$ [15].

Regarding the mechanisms behind this universal distribution, the most elementary firm dynamic model is that in Ref. [2], where Gibrat presented the law of proportionate growth to explain the lognormal distribution. However, by assuming

* Corresponding author. Tel.: +86 10 58802732; fax: +86 10 58210078.
E-mail address: qinghuachen@bnu.edu.cn (Q. Chen).

0378-4371/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2009.01.027

some version of a reflective lower bound, models of proportionate growth were widely applied to generate Pareto upper tail in economics [7,16–18]. Since firm size distribution provides another evidence of the universality of Zipf's law, the dynamics of firms gradually attracted the interests of physicists, especially those who work in econophysics [19–23]. Axtell insisted on using complexity approach to deal with the problem, and believed that agent-based modeling together with evolutionary dynamics should be helpful to understand the formation of power law [19]. Malcai et al. combined a random multiplicative dynamics at the individual level with a global coupling through a constraint given by the average of the agents' values to produce the power-law distribution [20]. Ree proposed an additive preferential rule to generate the power-law distribution [21]. Wang et al. constructed a bipartite producer-consumer network to explain the power-law distribution in firm size [22]. Reed and Hughes showed that if stochastic processes with exponential growth in expectation were killed randomly, it led to power-law tails in statistical distributions [23].

As listed above, many intriguing results have been achieved on firm size distribution, but there are still two aspects needing to be concerned further. Firstly, in contrast to these numerous works on developed economies, only a few researches have been carried out on developing countries [24]. Since some studies have indicated that company size distributions are not universal but clearly depend on countries [9,25]. It is very necessary to further widen the scope of this research. China, the biggest developing country in the world, attracted the eyes of the world for its market economy reform and sustaining high economic growth. Chinese companies have grown bigger and get more power in worldwide economic activities. For instance, *Sinopec* ranked 17th in *Fortune Global 500* of the year 2007, which only ranked 86th in 2002. For another example, *Lenovo*, has acquired *IBM*'s entire global desktop and notebook computer business, and gradually plays more and more important roles in world personal computer market these years. It is meaningful to check whether the Zipf's law holds up in such a unique ideological and political system. Secondly, although many theoretical models have been introduced to explain Zipf's law in firm size, most of them only use simple stochastic or physical mechanisms without sufficient economic senses. They ignore the real mechanism of economic operation and cannot give us any practical advices on the firm's real dynamics.

In this paper, we show the character of Chinese firms and try to give an economic explanation. The remainder is organized as follows: Section 2 carefully checks the data and displays the upper tail distribution of firms' sizes in China from the year 2002 to 2007. Section 3 has some discussions on it from the view of economics and in Section 4 we conclude.

2. The data and the distribution

The datasets used in this paper were published in series volumes of "*A Report on the Development of China's Enterprises*". They are provided annually by the organization named *China Enterprise Confederation* since the year 2002. The data consist of some basic information of top 500 Chinese enterprises ranked in revenue scale, including the firms' names, their revenues and corresponding ranks. Till now, six years' data are available.

It is very clear that the list of top 500 enterprises is changing, the revenue and corresponding rank for the same firm also fluctuate all the time. For example, *China National Heavy Duty Truck Group Corp.*, its revenue is respectively 5.159, 10.231, 15.124, 23.389, 14.394, 13.263 billion RMB, and its rank is 214, 143, 124, 107, 216, 289 from the year 2002 to 2007. Further more, we also find out (I) revenues of firms are mostly below 200 billion RMB, only a few firms are of great revenues while small firms are extremely common; (II) there is a big gap of revenue between the 1st firm and the 500th one, for instance, in 2007 the revenue of the 1st firm reaches 1064.667 billion RMB while that of the 500th only amounts to 7.216 billion RMB; (III) the gap between two contiguous firms ranking upper is much larger. As an example, the gaps between the 1st and 2nd, 249th and 250th, 499th and 500th are respectively 56 733.550, 40.150, 5.420 million RMB in the year 2002.

Fig. 1 shows the relation between revenues and their ranks of these firms in double logarithmic scale. Almost all of these curves parallel with the dashed straight lines with a slope of −1. As a result, the revenues of the top companies in China could be primarily concluded to exhibit Zipf's law. To get the precise Zipf's exponent, we use the fitting technique proposed by Gabaix et al. [26], which is a modified OLS estimation as below:

$$\log(S) = \alpha - \beta \log(r - 0.5), \quad (1)$$

where S is the firm size, r is its rank, and β is the Zipf's exponent needing to be estimated. This estimating can reduce the bias and performs well in the case of small samples where OLS procedure is strongly biased. The true standard error here, not the OLS standard error, is asymptotically $\sqrt{\frac{2}{n}}\beta$, where n is the number of samples.

In this case, the estimate and test are carried out by choosing segments in the rank list from 20 to 500 for all these six years. The results of estimate are listed in Table 1, including the estimates of exponent ($\hat{\beta}$), their corresponding determination coefficients (R^2) of these regression equations and the standard deviations (σ) defined above. It is shown that the exponent estimates of the six years are all around 1. Then we use the Kolmogorov–Smirnov (KS) statistics [27] to confirm the fitness of our result. The KS statistic is simply the maximum distance between the cumulative distribution functions (CDF) of the data and that of the fitted model:

$$D = \max_{x \in R} |F(x) - P(x)|\,, \quad (2)$$

where $F(x)$ denotes the empirical CDF and $P(x)$ is the hypothesized CDF. As listed in Table 1, the KS statistics are all less than 0.04. These results indicate that it is insufficient to reject the conclusion that the revenues of the samples follow a Zipf distribution.

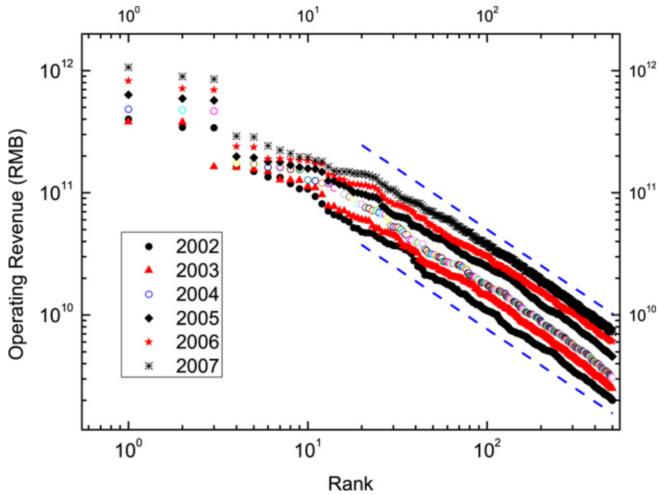

Fig. 1. The revenue distribution of top 500 Chinese enterprises in log–log scale. The dashed straight lines with a slope of −1.

Table 1
Estimates and KS statistics for the distribution of top 500 Chinese enterprises.

Year	$\hat{\beta}$	R^2	σ	KS statistics
2002	1.013	0.998	0.065	0.005
2003	1.006	0.996	0.065	0.021
2004	1.010	0.996	0.065	0.016
2005	0.971	0.995	0.063	0.034
2006	0.937	0.997	0.060	0.039
2007	0.960	0.997	0.062	0.036

3. Economic explanation and discussion

As mentioned in Section 1, most of theoretical models on firm size distribution are based on simple stochastic or physical mechanisms. But, for understanding this phenomenon, it is necessary to pay more attention to the reality of firm size formation. In an economic system, each firm possesses of a certain amount of capital and will employ labors to provide product or service under a certain production condition. The revenue can be realized while product or service is traded in the market. Then, this revenue will be mainly used for reward to production factors and investment to capital. The investment, along with the depreciation of previous capital, will form firm's capital stock. The firm may obtain more revenue under the production with more capital. Thus the distribution of revenues is mainly shaped by production technology and capital stock.

Actually for Chinese firms, the human resource is more abundant and much cheaper, so it doesn't play as important roles as other production factors. Capital and production technology are the determinants of the production process in China. As is known, AK model is usually used to describe how the combination of capital and technology produces output in the classical economics [28–30]. This model just represents the firm size formation process in China [31].

According to the analysis presented above, the ideal mechanism of the revenue formation can be explained as follows.

(1) The economy is characterized by a fixed number of firms N with the same initial value of capital K_0.

(2) At time t, the revenue of firm $iY_i(t)$ is determined by it's capital $K_i(t)$ and technology level $A_i(t)$ at this time. This process can be described as

$$Y_i(t) = A_i(t) \times K_i(t), \tag{3}$$

where $A_i(t)$ is assumed to be a random variable.

(3) Firm's capital spontaneously decreases due to physical or economic depreciation. In order to maintain or expand production, some share of the firm's previous revenue should be put into investment. At time t, the capital accumulation of firm i is given by:

$$K_i(t) = K_i(t-1) + s_i(t)\,Y_i(t-1) - \delta_i(t-1)\,K_i(t-1), \tag{4}$$

where δ is the depreciation rate, s is the share for investment.

(4) Here we consider the death or bankruptcy of firms and the entry of new firms by setting a minimum capital $K_{\min}$ in the economy. It means that if the capital of a firm is too small, it can not produce normally. When the capital is less than $K_{\min}$, we reset it to be $K_{\min}$. This procedure can be thought as an enforced identity of the death of an old company and the entry of a new firm.

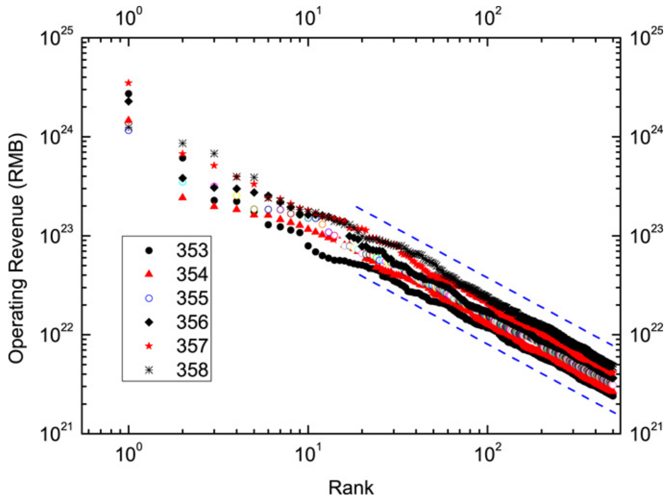

Fig. 2. The revenue distribution of the simulation result in log–log scale. The dashed straight lines with a slope of −1.

To check the validity of AK model on the firm growth, we have performed a numerical simulation. In the present work, the initial number of firms is set as $N = 1000\,000$ and the initial value of all firms' capital is assumed to be $K_0 = 0.01$. For simplicity, we set $K_{\min} = 0.01$. A, s and δ are all drawn from uniform distributions ($A, s \in [0, 1]$, $\delta \in [0, 0.2]$). The top 500 firms' revenues and ranks of the simulation result are plotted in Fig. 2. From this figure, we can draw the same conclusion that the distribution of the largest 500 revenues in our model follows Zipf's law.

Performing the simulation of this model, we reproduce the Zipf character of the top 500 Chinese companies. From the statistical view, the evidence on the recurrence of Zipf's law in our model is obvious. According to Eqs. (3) and (4), we can get

$$\begin{cases} K_i(t) = [1 + s_i(t)A_i(t-1) - \delta_i(t-1)]\,K_i(t-1) \\ Y_i(t) = \dfrac{A_i(t)}{A_i(t-1)}[1 + s_i(t)A_i(t-1) - \delta_i(t-1)]\,Y_i(t-1)\,. \end{cases} \tag{5}$$

Considering the constraint of the existence of $K_{\min}$, it is actually a multiplicative process model which has been proved to exhibit power law or Zipf's law at the upper tail [7].

Although our model degenerates to one kind of multiplicative process, it is incorporated with more economic senses. Here we argue that more economic details should be considered as we explore the mechanism of firm size distribution. A few similar attempts have been made recently [32,33].

4. Conclusions

In this paper, we have investigated the scaling behavior in Chinese firms. The statistical features of Chinese enterprises' revenues are presented in the framework of econophysics. By using the estimate technique of tail exponent provided by Gabaix and Ibragimov, we find that the sizes of top 500 Chinese enterprises follow Zipf's law with exponent of 1 in the recent six years, and this conclusion is confirmed by the KS test. This result is consistent with the cases in US, Japan and many other countries. Although most of the large companies are state-owned and the marketing is not fully dominant in the whole economy, some common characters for economic systems have emerged in China. In contrast to statistical physical model, in this work, we provide an economic explanation for this scaling behavior in China on the basis of AK model.

Acknowledgments

This research was supported by National Natural Science Foundation of China under Grants of No. 70771012 and No. 60534080.

References

[1] Y. Ijiri, H.A. Simon, Skew Distributions and the Sizes of Business Firms, North-Holland, New York, 1977.
[2] R. Gibrat, Les Inégalités Economiques, Sirey, Paris, 1931.
[3] B.H. Hall, The Journal of Industrial Economics 35 (1987) 583–606.
[4] M.H.R. Stanley, S.V. Buldyrev, S. Havlin, R.N. Mantegna, A. Salinger, H.E. Stanley, Economics Letters 49 (1995) 453–457.
[5] L.A.N. Amaral, S.V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, M.H.R. Stanley, Journal of Physics I France 7 (1997) 621–633.
[6] J. Voit, Advances in Complex Systems 4 (2001) 149–162.

[7] H.A. Simon, C.P. Bonini, American Economic Review 48 (1958) 607–617.
[8] E. Gaffeo, M. Gallegati, A. Palestrini, Physica A 324 (2003) 117–123.
[9] T. Kaizoji, H. Iyetomi, Y. Ikeda, Evolutionary and Institutional Economics Review 2 (2006) 183–198.
[10] G. Zipf, Human Behavior and the Principle of Last Effort, Addison-Wesley, Cambridge, MA, 1949.
[11] K. Okuyama, M. Takayasu, H. Takayasu, Physica A 269 (1999) 125–131.
[12] T. Mizuno, M. Katori, H. Takayasu, M. Takayasu, Statistical laws in the income of Japanese companies, in: H. Takayasu (Ed.), Empirical Science of Financial Fluctuations- The Advent of Econophysics, Springer Verlag, Tokyo, 2002, pp. 321–330.
[13] R.L. Axtell, Science 293 (2001) 1818–1820.
[14] Y. Fujiwara, H. Aoyama, C. Di Guilmi, W. Souma, M. Gallegati, Physica A 344 (2004) 112–116.
[15] M.E.J. Newman, Contemporary Physics 46 (2005) 323–351.
[16] J. Sutton, Journal of Economic Literature 35 (1997) 40–59.
[17] D.G. Champernowne, Economic Journal 63 (1953) 318–351.
[18] X. Gabaix, Quarterly Journal of Economics 114 (1999) 739–767.
[19] R.L. Axtell, The emergence of firms in a population of agents: Local increasing returns, unstable nash equilibria, and power law size distributions, CSED Working Paper No. 3, Brookings Institution, 1999.
[20] O. Malcai, O. Biham, S. Solomon, Physical Review E 60 (1999) 1299–1303.
[21] S. Ree, Physical Review E 73 (2006) 026115.
[22] D. Wang, L. Zhou, Z. Di, Physica A 363 (2006) 359–366.
[23] W.J. Reed, B.D. Hughes, Physical Review E 66 (2002) 067103.
[24] R. Hernández-Pérez, F. Angulo-Brown, D. Tun, Physica A 359 (2006) 607–618.
[25] H. Takayasu, K. Okuyama, Fractals 6 (1998) 67–79.
[26] X. Gabaix, R. Ibragimov, Rank-1/2: A simple way to improve the OLS estimation of tail exponents, NBER Technical Working Paper, 2007, p. 342. At http://www.nber.org/papers/t0342.pdf.
[27] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C: The Art of Scientific Computing, 2nd ed., Cambridge University Press, Cambridge, 1992.
[28] P. Romer, Journal of Political Economy 94 (1986) 1002–1037.
[29] R. Lucas, Journal of Monetary Economics 22 (1988) 3–42.
[30] R. Barro, Journal of Political Economy 99 (1991) 103–125.
[31] Y. Su, X. Xu, Economic Research Journal 11 (2002) 3–11.
[32] H. Li, Y. Gao, Physica A 387 (2008) 5225–5230.
[33] Q. Ma, Y. Chen, H. Tong, Z. Di, Physica A 387 (2008) 3209–3217.