Physica A 436 (2015) 933–942

Contents lists available at ScienceDirect

Physica A

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

An agent-based interaction model for Chinese personal income distribution

Yijiang Zou, Weibing Deng *, Wei Li *, Xu Cai

Complexity Science Center, Institute of Particle Physics, Central China Normal University, Wuhan 430079, China

HIGHLIGHTS

- The personal income distribution in China was studied by employing the empirical data.
- Results from an agent-based model with mean-field approach and the model on networks, the simulation and empirical data have been compared.
- Consider the high-order neighbors in networks as the global factor when agents exchange their wealth in this model.
- To characterize the empirical findings, simulation on different networks has been studied.

ARTICLE INFO

Article history:
Received 18 November 2014
Received in revised form 18 April 2015
Available online 23 May 2015

Keywords:
Chinese personal income distribution
Pareto law
Mean-field approach
Agent-based model

ABSTRACT

The personal income distribution in China was studied by employing the data from China Household Income Projects (CHIP) between 1990 and 2002. It was observed that the low and middle income regions could be described by the log-normal law, while the large income region could be well fitted by the power law. To characterize these empirical findings, a stochastic interactive model with mean-field approach was discussed, and the analytic result shows that the wealth distribution is of the Pareto type. Then we explored the agent-based model on networks, in which the exchange of wealth among agents depends on their connectivity. Numerical results suggest that the wealth of agents would largely rely on their connectivity, and the Pareto index of the simulated wealth distributions is comparable to those of the empirical data. The Pareto behavior of the tails of the empirical wealth distributions is consistent with that of the 'mean-field' model, as well as numerical simulations.

1. Introduction

During the past century, the human wealth distribution has been extensively studied [1]. In 1906, Italian economist Vilfredo Pareto [2] observed that 80% of the land in Italy was owned by 20% of the population, then he carried out such surveys on a variety of other domains, and found the similar rule also applied to people's personal income or wealth, company's profits/sales, *etc.* This is the well known Pareto law (or 80–20 rule). Due to the development of computer science and internet technology, there are more financial economic data available. Similar studies have been conducted in many countries, such as Australia [3], Brazil [4], India [5], China [6,7], Italy [8], Japan [9], France and Germany [10], the United Kingdom and the United States [11]. These works confirmed the universality of Pareto's law with much improved statistics. Further explorations revealed that this finding only applied to the top 1%–3% population with high income, while the majority (97%–99%) population with middle and low income obeyed log-normal law [12].

* Corresponding authors.
E-mail addresses: wdeng@mail.ccnu.edu.cn (W. Deng), liw@mail.ccnu.edu.cn (W. Li).

http://dx.doi.org/10.1016/j.physa.2015.05.086

Due to the recent rapid development of Chinese economy, several studies have been done based on the Chinese economical data. Zhou [13] performed a comparative study of the stock market and real-estate in China and some western countries. They found that the characteristic parameters of the anti-bubble exhibited remarkable stability over one year. Jiang [14] investigated the statistical properties of the empirical data taken from the Chinese stock market during the time period from January, 2006 to July, 2007. Deng [15] found fractal behavior and long-range correlations in the return series by testing 88 different funds of the Chinese fund market (CFM). A power-law relationship between the deviation D of prices and the Hurst exponent H was obtained. As to the Chinese personal income, some parallel studies have also been carried out. Guo [6] successfully applied the available survey data from China National Bureau of Statistics (NBS) to describe the individual income distribution in China. Xu [7] found that the income distribution changed a lot from its appearance in the early stage of Chinese reform and turned out to be consistent with that of complete market economics.

Other than the empirical analysis, many numerical models have been proposed to study wealth distribution from perspective of the stochastic dynamical process [16–19]. Bouchaud and Mézard [20] proposed a model in which the time evolution of wealth w_i of an agent i is given by a stochastic differential equation. Numeric result showed a power-law tail at large w and a sharp cutoff at small w. Slanina [21] formulated a model of wealth production and exchange, in which agents randomly interact in pairs. They found that the tail of the wealth distribution has a power-law form, and its exponent α is determined by the interplay between the intensity of the wealth exchange and the amount of wealth produced. Scafetta [22] presented a computational nonlinear stochastic model for the wealth distribution that depends upon two mechanisms: trade and investment.

However, to our best knowledge, there are few models that considered transactions in both higher-order neighbors and nearest neighbors in networks. So here we considered interactions with both two kinds of neighbors in the pair-interactive model, namely agents can trade with any other agents through their nearest neighbors. We consider the shortest path as the global trading path when trading happens by long-range interaction. The degree of node is regarded as the intensity of local effect when agents interact with their nearest neighbors. Meanwhile, to compare with the agent-based model on networks, a stochastic interactive model with the same exchange wealth among individuals was also studied. In this 'mean-field' model, all individuals have the same rate when they exchange the wealth with others.

This paper is organized as follows: in Section 2, the empirical data of the Chinese personal income was studied from various aspects. The stochastic interaction model was proposed and analyzed in Section 3, attempting to interpret the empirical findings. Section 4 presented the simulation results of our model on different networks. Conclusions were given in Section 5.

2. Income distribution in China

The China Household Income Projects (CHIP) collected the data of personal income from more than 15 000 rural households of 22 provinces in China, as well as more than 10 000 urban households from 1990 to 2002. As known, for urban households, the incomes are based on each individual, but for rural households, the incomes are based on the whole family. With the assumption that incomes are distributed equally among family members, the household income is transformed into the equivalent personal income [1,23]. The equivalence scale adopted to equivalize incomes for households of different sizes is $y^e_{it} = y_{it}/n^{0.66}_{it}$, where y^e_{it} is the equivalent income of individual i at time t. y_{it} and n_{it} are the total incomes and the size of the household to which individual i belongs at time t, respectively [24]. Therefore, the personal income data includes the income of both rural and urban individuals.

Firstly, we would like to uncover whether there are some patterns or common features embedded in the personal income distribution across different years. In order to reduce the statistical errors and effects of finite statistics, we studied the cumulative distribution function (CDF) of personal income, $P_>(x) = \int_x^\infty p(x)\mathrm{d}x$. The cumulative distributions of personal income in China from 1990 to 2002 are presented in Fig. 1.

We use the fitting metric – the maximum likelihood estimator (MLE) – proposed by Clauset et al. [25]. When the minimum point $x_{\min}$ is chosen, a more accurate power-law exponent α can be obtained. $x_{\min}$ is determined by minimizing the "distance" between the power law function and the empirical data, which is called Kolmogorov–Smirnov (KS) test:

$$D = \max_{x \ge x_{\min}} |S(x) - P(x)|, \tag{1}$$

here $S(x)$ is the CDF for the observation of the data with value at least $x_{\min}$, and $P(x)$ is the power-law model that best fits the data in the region $x \ge x_{\min}$, respectively. The estimated $x_{\min}$ is then the value of $x_{\min}$ that minimizes D. In the empirical data, we use $w_{\min}$ to represent the minimum income corresponding to the $x_{\min}$ in the KS test. So once the minimum point $w_{\min}$ of the empirical data is found, the high income part can be fitted with a power-law distribution. The slope of the straight line in the log–log figure is just the Pareto index α, which represents the scale of the distribution. From Fig. 1, one would find that in the low income range, it obeys the log-normal distribution $C(x) = 1/2\left[1 - \mathrm{Erf}(\frac{\log(x/m)}{s\sqrt{2}})\right]$, where $\mathrm{Erf}(x) = (\frac{2}{\sqrt{\pi}})\int_0^x e^{-z^2}\mathrm{d}z$ is the error function. This feature implies that the income gap is small, and the inequality is not very obvious in this region. However, the high income could be well fitted by the power law distribution $C(x) = (\frac{x}{x_{\min}})^{-\alpha+1}$. In this region, most wealth belong to a small number of people. Values of the fitting parameters of income distribution from the above figures are presented in Table 1. It is found that there are no regular changes in Pareto index, compared with the gradually increasing Gibrat index. The latter means that the income gap becomes smaller in the middle and low income region.