Available online at www.sciencedirect.com
ScienceDirect
Physica A 377 (2007) 230–240
PHYSICA A
www.elsevier.com/locate/physa

Income distribution: Boltzmann analysis and its extension

He Yuqing
Department of Administrative Management, Jinan University, Guangzhou 510632, Guangdong Province, P.R. China

Received 26 April 2006; received in revised form 3 November 2006
Available online 8 December 2006

Abstract

The paper aims at describing income distribution in moderate income regions. Starting with dividing income behaviors into the two parts: random and deterministic, and by introducing "instantaneous model" for theoretical derivations and "cumulative model" for positive tests, this paper applies the equilibrium approach of statistical mechanics in the study of nonconserved individual income course. The random income follows a stationary distribution similar to the Maxwell–Boltzmann distribution in the instantaneous model. Combining this result with marginal analysis, the probability distribution of individual income process that is composed of the random and deterministic income courses approximately obeys a distribution law mixing exponential function with a logarithmic prefactor. Using the census or income survey data of USA, UK, Japan, and New Zealand, the distribution law has been tested. The results show that it agrees very well with most of the empirical data. The discussion suggests that there might be essentially different income processes to happen in moderate and high income regions.

Keywords: Deterministic; Random; Income; Distribution

1. Introduction and general models

Since the pioneering power-law distribution of cumulative wealth above a level was proposed by Pareto in 1897 [1], many authors studied wealth and income distributions [2–36]. Diversity of the researching interests had evidently increased in the last decades. A number of studies concentrate on looking for empirical evidences for different distribution patterns [2–14], e.g. power-law, exponential, etc. Their results showed that the distribution patterns are different in low and high income tails, and power law is not a universal characteristic for individual income but suitable for the high income end (Reed held that the power law suits for low and high income tails but with different exponents [33]). In the wealth and income distributions of UK and USA, only about 5% of the total populations follow the power law in the high income tails [7]. Other studies introduced different models, by theoretical derivations or computer simulations, to interpret the basis of those empirical distributions (e.g. [14–36]). Many of these work mainly focused on providing explanations for higher- or lower-end power-law distributions (e.g. [15–26,29–31,33–35]), and some others primarily aimed at describing other distribution patterns (e.g. [14,21,27,28]). Our interest here belongs to the latter, and aims at

E-mail address: chnheyuqing@mail.china.com.

0378-4371/$ - see front matter
doi:10.1016/j.physa.2006.11.009

moderate but not high income region. The moderate income region contains most of population number, and thus has an essential importance.

It is well recognized that power law holds in the high end of individual income distribution. However, there is no commonly acceptable model for the moderate region of individual income distribution by now. Some theoretical models had been proposed involving in this region in literatures, such as exponential distribution [28], gamma distribution [14], log-normal distribution [18], double Pareto-lognormal distribution [33], various additive and multiplicative models (e.g. [21]), etc. No matter what theoretical ends the models aim at, they all have a common trait: they are based on thoroughly stochastic processes, such as random walk (e.g. [16]), random-pair exchange (e.g. [21,28]), random scattering course [35], geometric Brownian motion [33], Lotka–Volterra system [31], etc. It is however quite forced to modify all individual incomes into a thoroughly stochastic process in a market. Say, individual monthly salary, a most important income source for most of people which crucially determine the moderate income distribution, does not appear stochastically within a certain period, even does not evolve stochastically in relatively long a period (e.g. periodic step wage increase). The thorough randomness will be given up in this paper. Instead, a combination of partial randomness with nonrandomness will be adopted. This is the starting point of coming discussions, and the most outstanding characteristic distinguishing from other studies.

Equilibrium approaches in statistical physics usually relate to some conservation conditions. When econophysicists apply these approaches to describe economic systems, they have to restrict some economic quantities, such as circulating money, population, etc. in a rigorous conservation form unrealistically (e.g. [17,21,27,28]). Indeed, to apply the equilibrium model Boltzmann method in a realistic market, we must face the problem that the circulating money in an individual income market is increasingly cumulative with time. They compose of an open evolution system in which the economic quantities are nonconserved and nonequilibrium [37–39]. Computer simulations have shown that Boltzmann–Gibbs law applies to the description of random transactions between agents without saving propensity in a market with conserved monetary media [27,28]. Using the equilibrium approach of statistical mechanics in descriptions of nonconserved economic systems is another attempt of this paper.

The underlying models for solving the problem are based on the fact that within a properly limited piece of time, e.g. within a day, the amount of money actually circulating in a specific labor market can be viewed as a relatively fixed instantaneous quantity, and therefore can be treated as conserved approximately. Within such a limited piece of time, let n_i be the number of individuals on income level p_i, N the aggregate number of all individuals in the market, and M the aggregate income money of all individuals. The following relations will approximately hold:

$$\sum_i n_i = N \quad \text{(constant)}, \tag{1}$$

$$\sum_i n_i p_i = M \quad \text{(constant)}. \tag{2}$$

(1) and (2) are called "instantaneous model". Based on this model, the instantaneous quantity in a market is possibly studied in a framework of conservation theory approximately.

Let T stand for the evolution time. If a market is not very unrest, the cumulation will steadily evolve approximately. Corresponding to n_i, $n_i p_i$, N, and M in (1) and (2), the increasingly cumulative quantities observed in a real market are $n_i T$, $n_i p_i T$, NT, and MT approximately. They follow

$$\sum_i n_i T = NT, \tag{3}$$

$$\sum_i n_i p_i T = MT. \tag{4}$$

They compose of a nonconserved evolution system with time. (3) and (4) are called "cumulative model". If a cumulative quantity periodically, say seasonally evolves, the corresponding instantaneous quantity can be explained as average instantaneous quantity. Essentially the cumulative model is a linear approximation. If it is limited within not too long a period and the growing rate of income is not so high, the instantaneous and