Physica A 406 (2014) 155–162

Contents lists available at ScienceDirect

Physica A

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

Do wealth distributions follow power laws? Evidence from 'rich lists'

Michal Brzezinski

Faculty of Economic Sciences, University of Warsaw, Dluga 44/50, 00-241, Warsaw, Poland

HIGHLIGHTS

- We model the upper tail of wealth data taken from various 'rich lists'.
- Power-law model is consistent with data only in 35% of analysed data sets.
- Even if data are consistent with a power law, other distributions are plausible too.
- Most plausible alternatives include the log-normal and stretched exponential models.

ARTICLE INFO

Article history:
Received 31 March 2013
Received in revised form 22 October 2013
Available online 22 March 2014

Keywords:
Power-law model
Wealth distribution
Goodness-of-fit
Model selection
Pareto model

ABSTRACT

We use data on the wealth of the richest persons taken from the 'rich lists' provided by business magazines like *Forbes* to verify if the upper tails of wealth distributions follow, as often claimed, a power-law behaviour. The data sets used cover the world's richest persons over 1996–2012, the richest Americans over 1988–2012, the richest Chinese over 2006–2012, and the richest Russians over 2004–2011. Using a recently introduced comprehensive empirical methodology for detecting power laws, which allows for testing the goodness of fit as well as for comparing the power-law model with rival distributions, we find that a power-law model is consistent with data only in 35% of the analysed data sets. Moreover, even if wealth data are consistent with the power-law model, they are usually also consistent with some rivals like the log-normal or stretched exponential distributions.

1. Introduction

The search for universal regularities in income and wealth distributions has started over one hundred years ago with the famous work of Ref. [1]. His work suggested that the upper tails of income and wealth distributions follow a power law, which for a quantity x is defined as a probability distribution $p(x)$ proportional to $x^{-\alpha}$, with $\alpha > 0$ being a positive shape parameter known as the Pareto (or power-law) exponent. Pareto's claim has been extensively tested empirically as well as studied theoretically [2–6]. The emerging consensus in the empirical econophysics literature is that the bulk of income and wealth distributions seems to follow the log-normal or gamma distributions, while the upper tail follows the power-law distribution. Recent empirical studies found a power-law behaviour in the distribution of income in Australia [7,8], Germany [9], India [10], Italy [9,11,7], Japan [12,13], the UK [14,9,15], and the United States [14,9,16]. Another group of studies discovered a power-law structure of the upper tail of modern wealth distributions in China [17], France [18], India [10,19], Sweden [20], the UK [18,14,20,21], and the United States [22,18,20,23]. Surprisingly, analogous results were obtained for the wealth distribution of aristocratic families in mediaeval Hungary [24] and for the distribution of house areas in ancient Egypt [25].

E-mail address: mbrzezinski@wne.uw.edu.pl.

http://dx.doi.org/10.1016/j.physa.2014.03.052
0378-4371/

However, as shown recently by Ref. [26], detecting power laws in empirical data may be a difficult task. Most of the existing empirical studies exploit the fact that the power-law distribution follows a straight line on a log–log plot with the power-law exponent equal to the absolute slope of the fitted line. The existence of power-law behaviour is often confirmed visually using such a plot, while the exponent is estimated using linear regression. Such an approach suffers, however, from several drawbacks [27,26]. First, the estimates of the slope of the regression line may be very biased. This often happens because researchers assume that all available data follow a power-law model, while normally, it is so only for the upper-tail values above some lower bound. This lower bound on the power-law behaviour should therefore be estimated, if we want to obtain an unbiased estimate of the power-law exponent. Second, the standard R^2 statistic for the fitted regression line cannot be treated as a reliable goodness-of-fit test for the power-law behaviour. This is because distributions different from power-law distributions can approximate a power-law behaviour over a large part of the range of the variables under study resulting in a high value of R^2. Third, even if traditional methods succeed in verifying that a power-law model is a good fit to a given data set, it is still possible that some alternative model fits the data better. A complete empirical analysis would therefore require conducting a statistical comparison of the power-law model with some other candidate distributions.

Using a more refined methodology for measuring power-law behaviour, Ref. [26] has shown that the distribution of wealth among the richest Americans in 2003 as compiled in *Forbes*' annual US 'rich list' is not fitted well by a power-law model. Recently, Ref. [28] has tested formally for a power-law behaviour in Forbes' data on the wealth of the world's billionaires for the years 2000–2009. He has found that the Kolmogorov–Smirnov, Anderson–Darling and χ^2 goodness-of-fit tests all reject power-law behaviour for each of the data sets he used. However, Ref. [28] has tested if the whole range of observations in his data sets follow a power-law behaviour, while in fact this may apply only to a subset of the very largest observations. A more appropriate methodology for detecting a power-law distribution would therefore include a procedure for estimating a lower bound on the power-law behaviour.

The present paper uses a complete empirical methodology for detecting power laws introduced by Ref. [26] to verify if the upper tails of wealth distributions obey the power-law model or if some alternative model fits the data better. We estimate both the power-law exponent and the lower bound on the power-law behaviour. We also use goodness-of-fit tests and compare power-law fits with fits of alternative models. We analyse a large number of data sets on wealth distributions published annually by *Forbes* and other business magazines concerning wealth of (1) the richest persons in the world, (2) the richest Americans, (3) the richest Chinese, and (4) the richest Russians.

The paper is organized as follows. Section 2 presents the statistical framework used for measuring and analysing power-law behaviour in empirical data introduced by Ref. [26]. Section 3 describes our data sets drawn from the lists of the richest persons published by*Forbes* and other sources, while Section 4 provides the empirical analysis. Section 5 concludes the paper.

2. Statistical methods

In order to detect a power-law behaviour in wealth distributions, we use a toolbox proposed by Ref. [26]. A density of continuous power-law model is given by

$$p(x) = \frac{\alpha - 1}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-\alpha}. \tag{1}$$

The maximum likelihood estimator (MLE) of the power-law exponent, α, is

$$\hat{\alpha} = 1 + n \left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{\min}} \right], \tag{2}$$

where $x_i, i = 1, \ldots, n$ are independent observations such that $x_i \geqslant x_{\min}$. The lower bound on the power-law behaviour, $x_{\min}$, will be estimated using the following procedure. For each $x_i \geqslant x_{\min}$, we estimate the exponent using the MLE and then we compute the well-known Kolmogorov–Smirnov (KS) statistic for the data and the fitted model. The estimate $\hat{x}_{\min}$ is then chosen as a value of x_i for which the KS statistic is the smallest.[1] The standard errors for estimated parameters are computed with standard bootstrap methods with 10,000 replications.

The next step in measuring power laws involves testing the goodness of fit. A positive result of such a test allows us to conclude that a power-law model is consistent with a given data set. Following Ref. [26] again, we use a test based on a semi-parametric bootstrap approach. The procedure starts with fitting a power-law model to data using the MLE for α and the KS-based estimator for $x_{\min}$ and calculating a KS statistic for this fit, k. Next, we generate a large number of bootstrap data sets that follow the originally fitted power-law model above the estimated $x_{\min}$ and have the same non-power-law distribution as the original data set below $\hat{x}_{\min}$. Then, power-law models are fitted to each of the generated data sets using the same methods as for the original data set, and the KS statistics are calculated. The fraction of data sets for which their own KS statistic is larger than k is the p-value of the test. It represents a probability that the KS statistic computed for data drawn from the power-law model fitted to the original data is at least as large as k. The power-law hypothesis is rejected if

[1] The Kolmogorov–Smirnov statistic was also proposed by Ref. [27] as a goodness-of-fit test for the discrete power-law model, assuming, however, that the lower bound on power-law behaviour is known.