Journal of Economic Behavior & Organization 84 (2012) 364–381

Contents lists available at SciVerse ScienceDirect

Journal of Economic Behavior & Organization

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

The double power law in income distribution: Explanations and evidence☆

Alexis Akira Toda*

Department of Economics, Yale University, 28 Hillhouse Avenue, New Haven, CT 06511, United States

ARTICLE INFO

Article history:
Received 25 September 2011
Received in revised form 22 February 2012
Accepted 18 April 2012
Available online 27 April 2012

JEL classification:
C51
C65
D31
J3

Keywords:
Anderson–Darling test
Diffusion processes
Fokker–Planck equation
Kolmogorov–Smirnov test
Laplace distribution
Mincer equation
Power law

ABSTRACT

Conditional on education and experience, the distribution of personal labor income appears to be double Pareto, a distribution that obeys the power law in both the upper and lower tails. In particular, the error term of the classical Mincer equation appears to be Laplace, or double exponential. This "double power law" is not rejected by goodness-of-fit tests. I compare two diffusion processes (one mean-reverting, the other unit root) with a stationary double Pareto distribution as a model of income dynamics. The data favors the mean-reverting process for modeling income dynamics over the unit root process.

> "[The] absence of a satisfactory theory of the personal distribution of income and of a theoretical bridge connecting the functional distribution of income with the personal distribution is a major gap in modern economic theory." (Friedman, 1953, p. 277)

1. Introduction

The last decade saw a renewed interest in the power law behavior of economic and financial data (Gabaix, 1999, 2009). While previous investigators have mostly focused on the upper tail since the discovery of power laws in the size distribution of income and wealth by Pareto (1896a,b),[1] there are some recent evidences that the power law (with a positive power) also holds for the lower tail (Reed, 2003; Reed and Wu, 2008; Toda, 2011), a phenomenon predicted long ago by Champernowne

☆ This paper benefited from conversations with (in alphabetical order) Simone Alfarano, Joseph Altonji, Sylvain Barde, Truman F. Bewley, Donald Brown, Xiaohong Chen, Maximiliano Dvorkin, Mishael Milaković, Kohta Mori, Eduardo Souza Rodrigues, and Edward Vytlacil. I thank two anonymous referees for suggestions that greatly improved the paper. The financial support from the Cowles Foundation, the Nakajima Foundation, and Yale University is greatly acknowledged.

* Tel.: +1 203 432 3576.
E-mail address: alexisakira.toda@yale.edu
URL: http://pantheon.yale.edu/ at37/main.html.

[1] Pareto himself states that "the [power] law applies only to incomes a little above the minimum" (Pareto, 1897, footnote on p. 501).

0167-2681/$ – see front matter
http://dx.doi.org/10.1016/j.jebo.2012.04.012

(1953). This "double power law" behavior is interesting in its own right but has not yet captured much attention in the literature. This paper addresses two questions: (1) is the double power law behavior in the income distribution real?, and (2) if so, what is the mechanism that generates the double power law behavior? The first question concerns income distribution and the second income dynamics.

Although income distribution and income dynamics are closely connected as the income distribution is the cross-sectional distribution of income dynamics, these two topics have evolved somehow separately in the literature. For instance, most studies on income distribution deal with the unconditional distribution (hence do not control for individual characteristics),[2] whereas the income dynamics literature does control for individual characteristics but seems to pay little attention to the resulting income distribution, let alone the double power law behavior.[3]

The key to connect the two topics lies in the literature of human capital that starts with Mincer (1958), which has successfully related education and experience to (the conditional mean of) personal income (Lemieux, 2006). We focus on the error term of the classical Mincer equation (Mincer, 1974) to explain the double power law behavior of the conditional income distribution.

This paper consists of two parts. In Section 2, using cross-sections of U.S. male labor income data drawn from the Current Population Survey (CPS) and the Panel Study of Income Dynamics (PSID), I find by graphical inspection that the conditional income distribution (conditional on schooling and experience) is approximately double Pareto, or the conditional log income distribution is approximately Laplace. This is a much stronger claim than the traditional view that the upper tail of the income distribution obeys the power law (Pareto, 1896a; Mandelbrot, 1960). Although for our data the graphical test is convincing enough, I further check by goodness-of-fit tests because I believe that the failure of evaluating the income distribution models by specification tests[4] is partly responsible for the lack of consensus in the functional form of income distribution. Although the double Pareto conjecture (that the double Pareto distribution fits the conditional income distribution in the entire range) is rejected by the Kolmogorov test, the weaker conjecture that the conditional income distribution obey the double power law is not.

In Section 3, I compare two income dynamics models with stationary double Pareto distributions. The one proposed by Alfarano et al. (2012) in the context of the firm profit rates (returns on assets) is mean-reverting, whereas the other proposed independently by Gabaix (1999) and Reed (2001, 2003) is a unit root process. Using a panel of U.S. male labor income data for 1968–1993 drawn from the PSID, I find that the data favors the model of Alfarano et al. (2012).

2. Facts about conditional income distribution

In my earlier paper (Toda, 2011) I reported that the size distribution of income is *double Pareto* after controlling for individual trends, which were obtained by HP-filtering (Hodrick and Prescott, 1997) a panel data of income drawn from the PSID.[5] In this paper I go one step further and conjecture that such trends are given by the classical Mincer equation (Mincer, 1974)

$$\log Y = \beta_0 + \beta_1 S + \beta_2 E + \beta_3 E^2 + \epsilon, \tag{1}$$

where Y is personal income, S is years of school completed, and E is "experience" defined by $E = \text{Age} - S - 6$. In particular, the error term in the Mincer equation (1) appears to come from a *Laplace* distribution with mode at zero. Before presenting the evidence, I describe the basic properties of the power law and the double Pareto and Laplace distributions.

2.1. Digression: power law and double Pareto distribution

A nonnegative random variable X obeys the power law with exponent $\alpha > 0$ if $\lim_{x\to\infty} x^\alpha P(X > x) > 0$ exists (Pareto, 1896a, 1897; Mandelbrot, 1960, 1961, 1963).[6] We say that a nonnegative random variable X obeys the *power law in the lower tail* with exponent $\beta > 0$ if $\lim_{x\to 0} x^{-\beta} P(X < x) > 0$ exists, and X obeys the *double power law* with exponents α, β if it obeys the power law in both tails with exponents α, β, respectively. To the best of my knowledge, this paper is the first to introduce and define the concept of the double power law. It is easy to see that if X obeys the power law in the *lower* tail with exponent β, then $1/X$ obeys the power law in the *upper* tail with exponent β. Also, if X obeys the power law in the *upper* tail with exponent α, the kth moment $\mathrm{E}[X^k]$ does not exist if $k \geq \alpha$. In particular, the variance of X is finite if and only if $\alpha > 2$ and the mean of X exists if and only if $\alpha > 1$. Typical empirical results are that $\alpha \in (1.5, 3)$ for income and $\alpha \approx 1.5$ for wealth (Gabaix, 2008, 2009), hence the variance is infinite for wealth and may be infinite for income.

[2] For reviews of income distribution models, see Neal and Rosen (2000), Kleiber and Kotz (2003), and the articles collected in Chotikapanich (2008).

[3] See, for instance, Lillard and Willis (1978), Lillard and Weiss (1979), MaCurdy (1982), Meghir and Pistaferri (2004), Browning et al. (2010).

[4] A few exceptions that perform the Kolmogorov test are Dagum (1980) and Majumder and Chakravarty (1990).

[5] In this paper, the term "income" is generally used as an income measure, although all data are labor income (*i.e.*, earnings). The terms "income distribution" and "size distribution of income" are used synonymously.

[6] See Gabaix (2008) and Ibragimov (2009) for a quick introduction to the power law.