Physica A 433 (2015) 36–41

Contents lists available at ScienceDirect

Physica A

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

# Income distribution in the European Union versus in the United States

Maciej Jagielski *, Rafał Duczmal, Ryszard Kutner

*Faculty of Physics, University of Warsaw, Pasteura 5, PL-02093 Warszawa, Poland*

## HIGHLIGHTS

- We demonstrated the first comparison of income distribution in the EU and the US.
- Our formula describes the income of the EU households and the US individuals.
- Recent financial crisis was the most painful for the medium- and high-income classes.

## ARTICLE INFO

*Article history:*
Received 17 September 2014
Received in revised form 5 February 2015
Available online 2 April 2015

*Keywords:*
Income distribution
Yakovenko model
Financial crisis

## ABSTRACT

We prove that the refined approach – our extension of the Yakovenko et al. formalism – is universal in the sense that it describes well both household incomes in the European Union and individual incomes in the United States for all income social classes. This formalism, supplemented in this work by the entropy analysis, allowed the study of the impact of the recent world-wide financial crisis on the annual incomes of different income social classes. Hence, we find the most painful impact of the crisis on incomes of all income social classes. Furthermore, we indicate the existence of a possible market crisis precursor.



## 1. Introduction

One of the major trends having a long history in socio- and econophysics is the study of income and wealth redistribution in society, along with the analysis of social inequalities. Several models trying to explain the microscopic mechanisms of income dynamics of individuals or households were already proposed [1–28].

However, to the best of our knowledge, none of the models that have been developed so far give an analytic description of the annual household or individual incomes of *all* income social classes (i.e. the low-, medium-, and high-income classes) by a single formula based on a unified formalism. Recently [29], we extended the Yakovenko et al. model providing, indeed, such a unified formalism. This formalism has, in fact, a Markovian origin, thereby an analysis in terms of entropy is well grounded.

In the present paper we show that the formula which we derived within this unified formalism, containing a low number of free parameters, satisfactorily reproduces the empirical complementary cumulative distribution functions (CCDFs) both for the European Union (EU) and for the United States (US). The cumulative distribution function[1] is the main statistical tool used in this context, that is, the descriptive statistic technique is involved herein to analyse the data. Notably, this technique is supplemented herein by an analysis in terms of entropy, thus enabling access to thermodynamics, which makes our conclusions more convincing.

* Corresponding author. Tel.: +48 22 5532730; fax: +48 22 5532999.
*E-mail address:* zagielski@interia.pl (M. Jagielski).

[1] The complementary cumulative distribution function is the probability that the independent stochastic variable takes a value larger than some fixed one.

http://dx.doi.org/10.1016/j.physa.2015.03.071

## 2. Comments on formula

To describe the income of all social classes in the US and the EU, we used our extended Yakovenko et al. formalism (EYF) [29,30].

As for the Yakovenko et al. model, the coexistence of multiplicative and additive processes on the level of the Langevin equation and hence the Fokker–Planck one, is also allowed for the EYF. That is, we assume that household or individual incomes are determined by: (i) systematic wages and salaries and/or (ii) random profits that go to households or individuals mainly through financial investments and/or capital gains. Furthermore, for the EYF we assume that the formalism of the income change is the same for the entire society, however, its detailed dynamics distinguishes well the ranges of individual income social classes, in particular, of the high-income social class from that of others (see Ref. [29] for details). As the Markovian master equation is well approximated by the Fokker–Planck equation [31], it allows access to thermodynamics by using conditional entropies (that is, entropies related to each income social class separately)—this approach is developed in Section 4.2 making our analysis more versatile.

Within the framework of the EYF, we found the equilibrium probability distribution function in the form [29,30]

$$P_{\mathrm{eq}}(m) \propto \begin{cases} \dfrac{\exp\left(-(m_0/T)\arctan(m/m_0)\right)}{[1+(m/m_0)^2]^{(\alpha+1)/2}}, & \text{if } m < m_1 \\ \dfrac{\exp\left(-(m_0/T_1)\arctan(m/m_0)\right)}{[1+(m/m_0)^2]^{(\alpha_1+1)/2}}, & \text{if } m \geq m_1 \end{cases} \tag{1}$$

where the $m_0$ parameter is a crossover (border) income between the low- and medium-income social classes, while the $m_1$ parameter is an analogous border income but between the medium- and high-income social classes. The $T$ parameter can be interpreted as an average income per household or individual within the low- and medium-income social classes. Meanwhile, interpretation of the $T_1$ parameter is given further in the text. The shape exponents $\alpha$ and $\alpha_1$ are the Pareto ones, describing the income inequality within the medium- and high-income social classes, respectively. A detailed description of the parameters' specification and estimation can be found in our previous paper [29]. The CCDF considered below is an integrated quantity of the above given distribution function.

## 3. Remarks on databases

In the case of the European Union, we exploit the empirical data from Eurostat's Survey on Income and Living Conditions (EU-SILC) [32–37] for the years 2005–2010. This database contains information on the demographic characteristics of households, their living conditions, as well as income as economic activity. In our analysis we chose the *total household gross income* variable. However, Eurostat's EU-SILC database contains only a few observations concerning the income of households belonging to the high-income social class, which is insufficient to be subjected to any statistical description. In order to improve the statistics for the high-income social class, we additionally analysed the effective income of billionaires[2] in the EU by using the Forbes ranking 'The World's Billionaires' [38] (see Refs. [29,30] for more details).

In the case of the United States we used the empirical data for the years 2005–2010 taken from the Internal Revenue Service (IRS)—the US government tax agency [39]. We chose the *adjusted gross income* variable as the only one accessible in the context of our comparative analysis. Similarly, as for the EU-SILC, the IRS database does not contain observations on individuals belonging to the high-income social class. Again, in order to consider the high-income social class, we additionally analysed the effective income of billionaires in the US by using the same Forbes ranking as mentioned above.

By using the EU-SILC database as well as a ranking of the richest Europeans and the IRS data set and a ranking of the richest Americans, we were able to consider incomes of all social classes thanks to the joint procedure presented in details in Refs. [29,30]. Thus, we obtained a data record sufficiently large for the statistical consideration of all social classes.

## 4. Results and discussion

### 4.1. Analysis by using CCDF

We compared the theoretical CCDF, based on the probability distribution function $P_{\mathrm{eq}}(m)$ given by Eq. (1), with: (i) the empirical CCDF of the annual total gross income of households in the EU and (ii) the corresponding empirical CCDF of the annual adjusted gross income of individuals in the US. During our research, we analysed the empirical CCDF constructed by using the well-known Weibull rank formula [40,41].

The resulting plots, each consisting of the theoretical (solid curves) and empirical (small circles) CCDFs for the EU (the EU curves) and for the US (the US curves), are presented in Figs. 1 and 2 for a typical year, 2007, and an exceptional year, 2009. By comparing the EU curves (shown in both figures) we prove that in the latter year the household and individual incomes were strongly affected by the recent worldwide financial crisis (which began around 2006–2007). It is commonly known that this crisis threatened to bring about the collapse of large financial institutions, which was prevented by the bailout of banks by national governments. In addition, it had a negative impact on the stock and housing markets, resulting

[2] The term *billionaire* used herein is equivalent (as in the US terminology) to the term *multimillionaire* used in the European terminology.