

Review

The Kaniadakis Distribution for the Analysis of Income and Wealth Data

Fabio Clementi 

Department of Political Science, Communication and International Relations, University of Macerata,
Via Don Minzoni 22/A, 62100 Macerata, Italy; fabio.clementi@unimc.it; Tel.: +39-0733-258-2560

Abstract: The paper reviews the “ κ -generalized distribution”, a statistical model for the analysis of income data. Basic analytical properties, interrelationships with other distributions, and standard measures of inequality such as the Gini index and the Lorenz curve are covered. An extension of the basic model that best fits wealth data is also discussed. The new and old empirical evidence presented in the article shows that the κ -generalized model of income/wealth is often in very good agreement with the observed data.

Keywords: income and wealth distribution; parametric modeling; κ -generalized model

1. Introduction

The past two decades have seen a resurgence of interest in the study of income and wealth distribution in both the physics [1–4] and economics [5–9] communities. Scholars have focused particularly on the empirical analysis of large data sets to infer the shape of income and wealth distributions and to develop theoretical models that can reproduce them.

Pareto’s observation that the number of people in a population whose income exceeds x is often well approximated by $Cx^{-\alpha}$ was a natural starting point for this field of analysis [10–13]. However, empirical research has shown that the Pareto distribution accurately models only high income levels, while it does a poor job of describing the lower end of distributions.

As research has continued, new models have been proposed to better describe the data, using either a combination of known statistical distributions [14–22] or parametric functional forms for the distribution as a whole. Among these, the two-parameter lognormal [23] and gamma [24] distributions were proposed as models for the size distributions of income and wealth, but later evidence showed that these models tend to exaggerate skewness and perform poorly at the upper end of the empirical distributions [25–28]. Three-parameter models such as the generalized gamma [29–32], Singh–Maddala [33], and Dagum Type I [34] provide better fits. These models converge to the Pareto model for large values of income/wealth and accurately describe lower and middle ranges.

Finally, models with more than three parameters have also been suggested to fit income and wealth data. For example, the generalized beta distribution of the second kind (GB2) is a four-parameter distribution that was first described by [35]. It fits the data very well and also includes some of the two- and three-parameter models mentioned above as special or limiting cases. (The generalized beta distribution of the first kind (GB1) [35] and the double Pareto-lognormal distribution [36] are other four-parameter models that fit the data well. Ref. [37] also developed the five-parameter generalized beta distribution family, which includes the GB1 and GB2 as special cases and all of the two- and three-parameter distributions nested inside them. In turn, the double Pareto-lognormal distribution has been generalized into a five-parameter family of distributions called the generalized double Pareto-lognormal distribution [38]. However, closed-form expressions for probability density and/or cumulative distribution functions do not always exist for



Citation: Clementi, F. The Kaniadakis Distribution for the Analysis of Income and Wealth Data. *Entropy* **2023**, *25*, 1141. <https://doi.org/10.3390/e25081141>

Academic Editors: Antonio M. Scarfone, Dionissios T. Hristopoulos and Sergio Luiz E. F. da Silva

Received: 27 June 2023

Revised: 27 July 2023

Accepted: 28 July 2023

Published: 30 July 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

these “super” models, making fitting them to data computationally difficult and slow due to the need to use numerical methods [39,40]).

Among models that seek to provide a unified framework for describing real-world data, including the power-law tails found in empirical distributions of income and wealth, the κ -generalized distribution has demonstrated exceptional performance and is often seen as a better alternative to other widely used parametric models. This model, which was initially introduced in 2007 and progressively expanded in the years that followed, has its origins in the framework of κ -generalized statistical mechanics [41–46]. It has a bulk very similar to the Weibull distribution and an upper tail that decays according to a Pareto power law for high values of income and wealth, providing a sort of middle ground between the two descriptions.

The purpose of this paper is to provide a comprehensive overview of the important results concerning the κ -generalized distribution. The desire to celebrate the 20th anniversary of Kaniadakis’ notable contribution and the belief that an interdisciplinary approach integrating statistical mechanics and economics may give novel insights into economic relationships motivated this work. Giorgio Kaniadakis played a pivotal role in the development of the κ -generalized model, making valuable and direct contributions to its conception. The intention behind presenting information on the fundamental statistical properties and empirical plausibility of this distribution is to convince the reader of its importance and usefulness for future exploration.

The paper is structured as follows. Section 2 introduces the κ -generalized model, covering topics such as interrelations with other distributions, basic statistical properties, and inferential aspects. Section 3 presents recent results of fitting the κ -generalized distribution to empirical income data corresponding to the distribution of household incomes in Greece, and compares the relative merits of alternative income size distribution models using the same data. Section 4 reviews empirical applications showing that the κ -generalized model is often in excellent agreement with observed income data; the κ -generalized mixture model for net worth distribution, which best fits wealth data, is also discussed in this section. Section 5 concludes the paper with some remarks.

2. The κ -Generalized Model for Income Distribution

The κ -generalized statistical model, named after [47], is based on the use of κ -deformed exponential and logarithmic functions introduced by Kaniadakis [41–43] in the context of special relativity. Within this framework, the ordinary exponential function $\exp(x)$ deforms into the generalized exponential function $\exp_{\kappa}(x)$ given by:

$$\exp_{\kappa}(x) = \left(\sqrt{1 + \kappa^2 x^2} + \kappa x \right)^{\frac{1}{\kappa}}, \quad x \in \mathbb{R}, \quad \kappa \in [0, 1). \quad (1)$$

The deformed logarithmic function $\ln_{\kappa}(x)$, which is defined as the inverse of (1), can be written as:

$$\ln_{\kappa}(x) = \frac{x^{\kappa} - x^{-\kappa}}{2\kappa}, \quad x \in \mathbb{R}_{+}. \quad (2)$$

Kaniadakis’ deformed functions have also been successfully used to analyze nonphysical systems. In economics, the κ -deformation has been used to study differentiated product markets [48,49], finance [50–55], and the distribution of income by size [47,56–62]. In the latter case, it is interesting to use such deformed functions because they can be used to statistically describe the entire spectrum of incomes, from the low to the middle range and up to the Pareto tail.

2.1. Definitions and Basic Properties

A random variable X is said to have a κ -generalized distribution, and we write $X \sim \kappa\text{-gen}(\alpha, \beta, \kappa)$, if it has a probability density function (PDF) given by:

$$f(x; \alpha, \beta, \kappa) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \frac{\exp_{\kappa}[-(x/\beta)^{\alpha}]}{\sqrt{1 + \kappa^2(x/\beta)^{2\alpha}}}, \quad x > 0, \quad \alpha, \beta > 0, \quad \kappa \in [0, 1). \quad (3)$$

Its cumulative distribution function (CDF) can be expressed as:

$$F(x; \alpha, \beta, \kappa) = 1 - \exp_{\kappa}[-(x/\beta)^{\alpha}]. \quad (4)$$

(For a complete description of the κ -generalized distributional properties, the reader is referred to [60] and the references cited therein. A heuristic derivation of the κ -generalized density, showing how this probability distribution emerges naturally within the field of κ -deformed analysis, is given in [61,63]).

Figure 1 illustrates the behavior of the κ -generalized PDF and the complementary CDF, $1 - F(x; \alpha, \beta, \kappa)$, for various parameter values.

Each of the three graph pairs holds two parameters constant and varies the remaining one.

The constant β is a characteristic scale that has the same dimension as income. For this reason, it takes into account the monetary unit and can be used to adjust for inflation and facilitate cross-country comparisons of income distributions expressed in different monetary units. Increases in the monetary unit result in a global increase in individual income and average income.

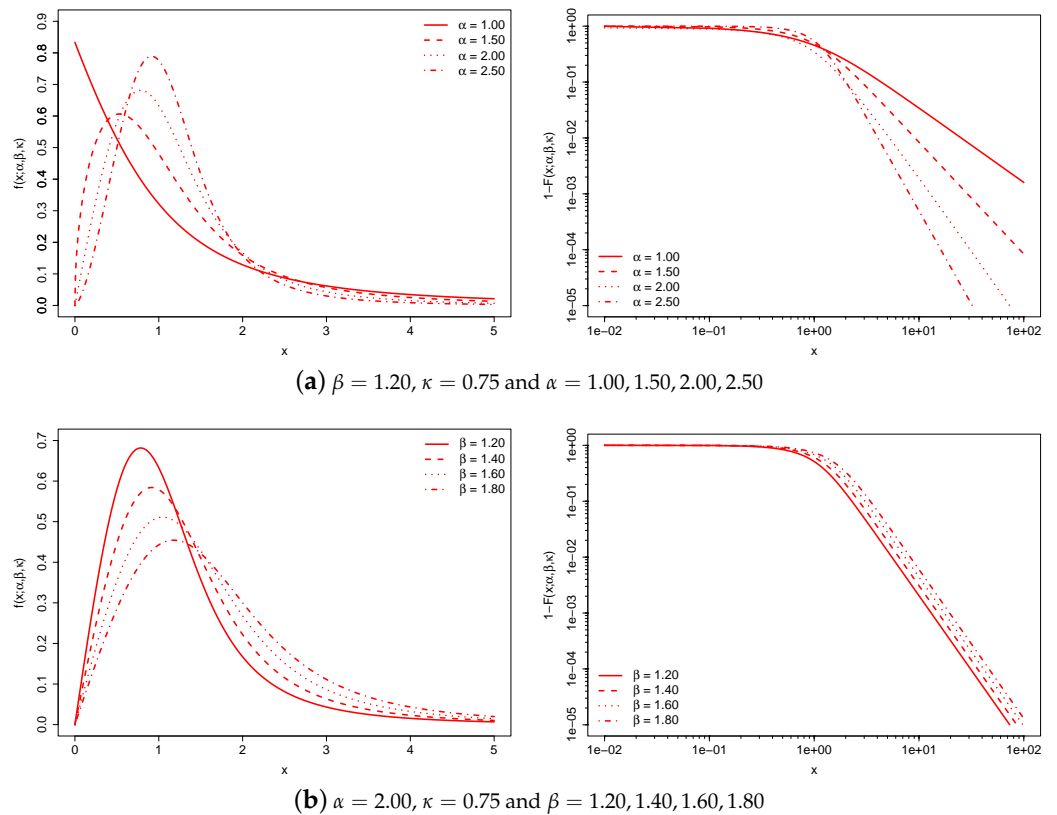
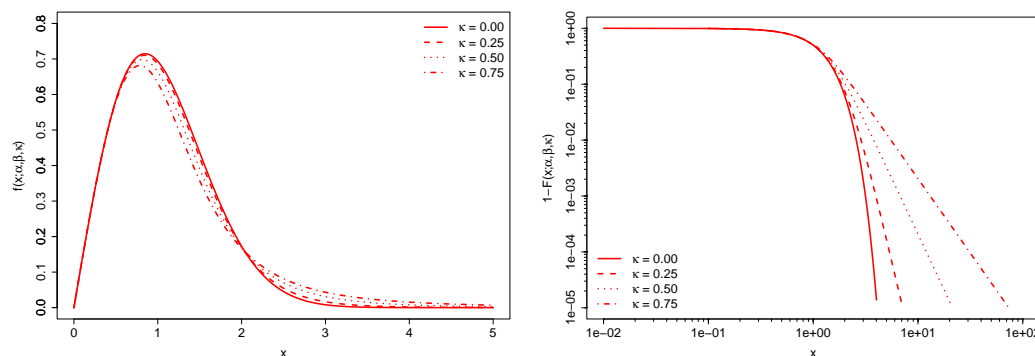


Figure 1. Cont.



(c) $\alpha = 2.00, \beta = 1.20$ and $\kappa = 0.00, 0.25, 0.50, 0.75$

Figure 1. κ -generalized PDF (left) and complementary CDF (right) for different values of the parameters. The complementary CDF is plotted on double-log axes, which is the standard way to emphasize the right-tail behavior of a distribution.

The α and κ parameters are scale-free parameters that affect the distribution’s shape. The region around the origin of the κ -generalized distribution is dominated by α , while the upper tail is dominated by both α and κ . Increasing κ leads to a thicker upper tail, while increasing α tapers both tails and increases the concentration of probability mass around the peak of the distribution.

As κ approaches 0, the distribution converges to the Weibull distribution; it is easy to verify that:

$$\lim_{\kappa \rightarrow 0} f(x; \alpha, \beta, \kappa) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp[-(x/\beta)^\alpha] \tag{5}$$

and:

$$\lim_{\kappa \rightarrow 0} F(x; \alpha, \beta, \kappa) = 1 - \exp[-(x/\beta)^\alpha]. \tag{6}$$

(The Weibull distribution is primarily studied in the engineering literature. In physics, it is known as the stretched exponential distribution when $\alpha < 1$. In economics, it has potential for income data, although it has only been used sporadically—some applications can be found in Refs. [29,35,64–69].) The distribution behaves similarly to the Weibull model for $x \rightarrow 0^+$, while for large x it approaches a Pareto distribution of the first kind with scale $k = \beta(2\kappa)^{-\frac{1}{\alpha}}$ and shape $a = \frac{\alpha}{\kappa}$, i.e.:

$$f(x; \alpha, \beta, \kappa) \underset{x \rightarrow +\infty}{\sim} \frac{ak^a}{x^{a+1}} \tag{7}$$

and:

$$F(x; \alpha, \beta, \kappa) \underset{x \rightarrow +\infty}{\sim} 1 - \left(\frac{k}{x}\right)^a, \tag{8}$$

thus satisfying the weak Pareto law [70]. (Additional versions of the Pareto law were introduced by [71], $\lim_{x \rightarrow +\infty} \frac{xf(x)}{1-F(x)} = a$, and [30], $\lim_{x \rightarrow +\infty} \left[1 + \frac{xf'(x)}{f(x)}\right] = -a$. Since we have: $\lim_{x \rightarrow +\infty} \frac{xf(x; \alpha, \beta, \kappa)}{1-F(x; \alpha, \beta, \kappa)} = \frac{\alpha}{\kappa} = a$ and $\lim_{x \rightarrow +\infty} \left[1 + \frac{xf'(x; \alpha, \beta, \kappa)}{f(x; \alpha, \beta, \kappa)}\right] = -\frac{\alpha}{\kappa} = -a$, the κ -generalized distribution also obeys these alternative versions of the weak Pareto law.)

Equation (4) implies that the quantile function is available in closed form:

$$F^{-1}(u; \alpha, \beta, \kappa) = x_u = \beta \left[\ln_\kappa \left(\frac{1}{1-u} \right) \right]^{\frac{1}{\alpha}}, \quad 0 < u < 1, \tag{9}$$

an attractive feature for generating random numbers from a κ -generalized distribution. The median of the distribution is:

$$x_{\text{med}} = \beta[\ln_{\kappa}(2)]^{\frac{1}{\alpha}}, \tag{10}$$

and the mode occurs at:

$$x_{\text{mode}} = \beta \left[\frac{\alpha^2 + 2\kappa^2(\alpha - 1)}{2\kappa^2(\alpha^2 - \kappa^2)} \right]^{\frac{1}{2\alpha}} \left\{ \sqrt{1 + \frac{4\kappa^2(\alpha^2 - \kappa^2)(\alpha - 1)^2}{[\alpha^2 + 2\kappa^2(\alpha - 1)]^2}} - 1 \right\}^{\frac{1}{2\alpha}} \tag{11}$$

if $\alpha > 1$; otherwise, the distribution is zero-modal with a pole at the origin.

Finally, the r th raw moment of the κ -generalized distribution is equal to:

$$\langle x^r \rangle = \int_0^{\infty} x^r f(x; \alpha, \beta, \kappa) dx = \beta^r (2\kappa)^{-\frac{r}{\alpha}} \frac{\Gamma(1 + \frac{r}{\alpha}) \Gamma(\frac{1}{2\kappa} - \frac{r}{2\alpha})}{1 + \frac{r}{\alpha}\kappa \Gamma(\frac{1}{2\kappa} + \frac{r}{2\alpha})}, \tag{12}$$

where $\Gamma(\cdot)$ denotes the gamma function, and exists for $-\alpha < r < \frac{\alpha}{\kappa}$. Specifically:

$$\langle x \rangle = \beta(2\kappa)^{-\frac{1}{\alpha}} \frac{\Gamma(1 + \frac{1}{\alpha}) \Gamma(\frac{1}{2\kappa} - \frac{1}{2\alpha})}{1 + \frac{1}{\alpha}\kappa \Gamma(\frac{1}{2\kappa} + \frac{1}{2\alpha})} \tag{13}$$

is the mean of the distribution and:

$$\langle x^2 \rangle - \langle x \rangle^2 = \beta^2 (2\kappa)^{-\frac{2}{\alpha}} \left\{ \frac{\Gamma(1 + \frac{2}{\alpha}) \Gamma(\frac{1}{2\kappa} - \frac{1}{\alpha})}{1 + \frac{2}{\alpha}\kappa \Gamma(\frac{1}{2\kappa} + \frac{1}{\alpha})} - \left[\frac{\Gamma(1 + \frac{1}{\alpha}) \Gamma(\frac{1}{2\kappa} - \frac{1}{2\alpha})}{1 + \frac{\kappa}{\alpha} \Gamma(\frac{1}{2\kappa} + \frac{1}{2\alpha})} \right]^2 \right\} \tag{14}$$

is the variance.

2.2. Measuring Income Inequality Using the κ -Generalized Distribution

The concept of inequality in economics dates back to Pareto’s early work [10–13], which showed that the top 20% of population held about 80% of total income/wealth. Later, the American economist Lorenz [72] introduced the Lorenz curve, a widely used tool for measuring income/wealth inequality. This curve measures the difference in income or wealth distribution from an equal distribution. If there is perfect equality, the Lorenz curve coincides with the diagonal of a unit square, while worsening distribution (more inequality) moves the curve away from the diagonal.

The Lorenz curve for a random variable X with CDF $F(x)$ and finite mean $\langle x \rangle = \int x dF(x)$ is defined as [73]:

$$L(u) = \frac{1}{\langle x \rangle} \int_0^u F^{-1}(t) dt, \quad u \in [0, 1]. \tag{15}$$

Using the closed form of the quantile function $F^{-1}(u)$ of the κ -generalized distribution, the Lorenz curve can be reformulated as follows [74]:

$$L(u) = I_x \left(1 + \frac{1}{\alpha}, \frac{1}{2\kappa} - \frac{1}{2\alpha} \right), \quad x = 1 - (1 - u)^{2\kappa}, \tag{16}$$

where $I_x(\cdot, \cdot)$ is the regularized incomplete beta function defined in terms of the incomplete beta function and the complete beta function, that is, $I_x(\cdot, \cdot) = \frac{B_x(\cdot, \cdot)}{B(\cdot, \cdot)}$. The curve (16) exists if and only if $\frac{\alpha}{\kappa} > 1$. In particular, if $X_i \sim \kappa\text{-gen}(\alpha_i, \beta_i, \kappa_i)$, $i = 1, 2$, the necessary and

sufficient conditions for the Lorenz curves of X_1 and X_2 not to intersect (otherwise, it would be impossible to determine which distribution has more inequality) are [58]:

$$\alpha_1 \geq \alpha_2 \quad \text{and} \quad \frac{\alpha_1}{\kappa_1} \geq \frac{\alpha_2}{\kappa_2}. \tag{17}$$

The Lorenz curves of two κ -generalized distributions X_1 and X_2 with parameters chosen according to (17) are illustrated in Figure 2. The depicted curves indicate that X_1 exhibits lower inequality compared to X_2 , as the Lorenz curve of X_1 does not intersect or fall below that of X_2 .

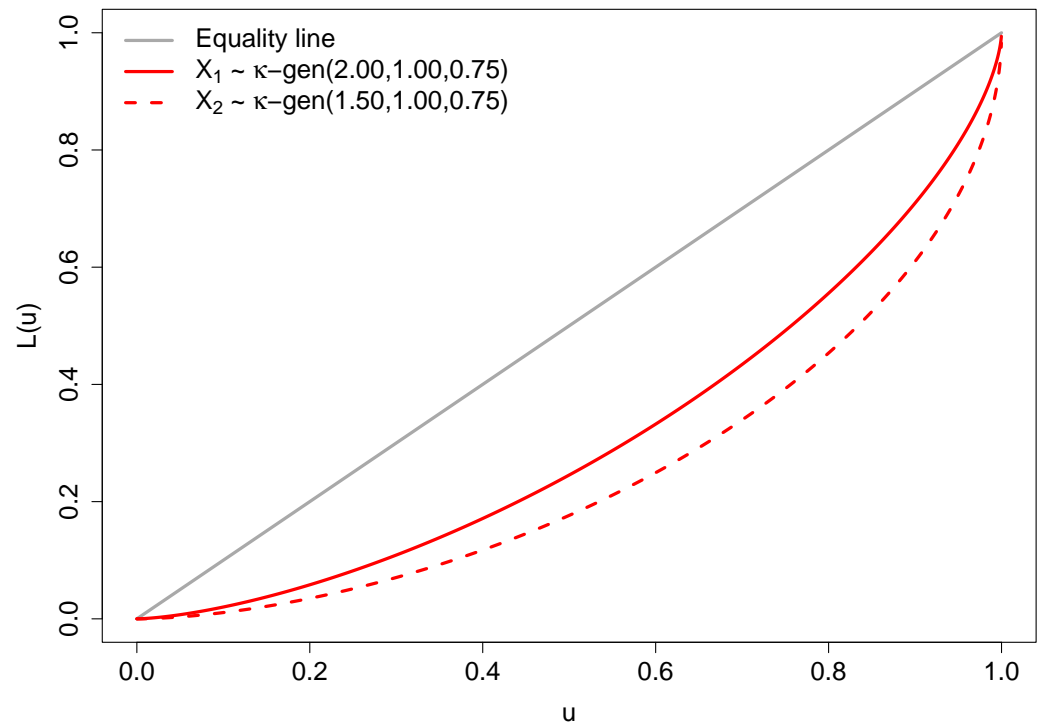


Figure 2. Lorenz curves for two κ -generalized distributions.

Economists have employed statistical metrics to quantify income and wealth inequality. The Gini coefficient, developed in 1914 by the Italian statistician Gini [75], is one of the best known. From the general definition $G = 1 - \frac{1}{\langle x \rangle} \int_0^\infty [1 - F(x)]^2 dx$ due to [76], the Gini coefficient associated with the κ -generalized distribution is:

$$G = 1 - \frac{2\alpha + 2\kappa}{2\alpha + \kappa} \frac{\Gamma\left(\frac{1}{\kappa} - \frac{1}{2\alpha}\right) \Gamma\left(\frac{1}{2\kappa} + \frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{\kappa} + \frac{1}{2\alpha}\right) \Gamma\left(\frac{1}{2\kappa} - \frac{1}{2\alpha}\right)}. \tag{18}$$

Using the Stirling approximation for the gamma function, $\Gamma(z) \approx \sqrt{2\pi}z^{z-\frac{1}{2}} \exp(-z)$, and taking the limit as $\kappa \rightarrow 0$ in Equation (18), after some simplification one arrives at $G = 1 - 2^{-\frac{1}{\alpha}}$, which is the explicit form of the Gini coefficient for the Weibull distribution (see e.g., [77], p. 177). Since the exponential distribution is a special case of the Weibull distribution with a shape parameter of 1, it follows directly that for $\kappa \rightarrow 0$ and $\alpha = 1$, the exponential law is also a special limiting case of the κ -generalized distribution with a true Gini coefficient of one half [16].

The Gini coefficient is a widely used measure of inequality, but it makes specific assumptions about income differences in different parts of the distribution. It is most sensitive to transfers around the middle of the income distribution and least sensitive to transfers among the very rich or very poor [78]. Differently, the generalized entropy class

of inequality measures [79–83] provides a range of bottom-to-top sensitive indices used by analysts to assess inequality in different parts of the income distribution. The expression for this class of inequality indices in terms of the κ -generalized parameters is [57]:

$$GE(\theta) = \frac{1}{\theta^2 - \theta} \left\{ \left(\frac{\beta}{m} \right)^\theta \left[\frac{(2\kappa)^{-\frac{\theta}{\alpha}} \Gamma\left(\frac{1}{2\kappa} - \frac{\theta}{2\alpha}\right)}{1 + \frac{\theta}{\alpha}\kappa \Gamma\left(\frac{1}{2\kappa} + \frac{\theta}{2\alpha}\right)} \Gamma\left(1 + \frac{\theta}{\alpha}\right) \right] - 1 \right\}, \quad \theta \neq 0, 1, \quad (19)$$

where $m = \langle x \rangle$ denotes the mean of the distribution given by Equation (13). Formula (19) defines a class because $GE(\theta)$ takes different forms depending on the value given to θ , the parameter that describes the sensitivity of the index to income differences in different parts of the income distribution—the more positive or negative θ is, the more sensitive $GE(\theta)$ is to income differences at the top or bottom of the distribution. Two limiting cases of (19), obtained when the parameter θ is set to 0 and 1, have gained attention in practical work for the purpose of measuring inequality; these are the mean logarithmic deviation index:

$$MLD = \lim_{\theta \rightarrow 0} GE(\theta) = \frac{1}{\alpha} \left[\gamma + \psi\left(\frac{1}{2\kappa}\right) + \ln(2\kappa) - \alpha \ln\left(\frac{\beta}{m}\right) + \kappa \right], \quad (20)$$

where $\gamma = -\psi(1)$ is the Euler–Mascheroni constant and $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the digamma function, and the Theil index [84]:

$$T = \lim_{\theta \rightarrow 1} GE(\theta) = \frac{1}{\alpha} \left[\psi\left(1 + \frac{1}{\alpha}\right) - \frac{1}{2}\psi\left(\frac{1}{2\kappa} - \frac{1}{2\alpha}\right) - \frac{1}{2}\psi\left(\frac{1}{2\kappa} + \frac{1}{2\alpha}\right) - \ln(2\kappa) + \alpha \ln\left(\frac{\beta}{m}\right) - \frac{\alpha\kappa}{\alpha + \kappa} \right], \quad (21)$$

where the former is more sensitive to variations in the lower tail, while the latter is more sensitive to variations in the upper tail [85]. (Equation (19) is not defined for $\theta = 0$ and $\theta = 1$, as $(\theta^2 - \theta) = 0$ in both cases. Expressions for these values of θ are therefore derived using l’Hôpital rule, which allows evaluating limits of indeterminate forms using derivatives. Expressions for any $GE(\theta)$ index other than the cases $\theta = 0, 1$ can be derived by simple substitution—see for example [60]).

Finally, the class of inequality measures introduced by Atkinson [86] can be derived from (19) by exploiting the relationship [87,88]:

$$A(\epsilon) = 1 - [\epsilon(\epsilon - 1)GE(1 - \epsilon) + 1]^{\frac{1}{1-\epsilon}}, \quad \epsilon > 0, \quad \epsilon \neq 1, \quad (22)$$

where $\epsilon = 1 - \theta$ is the inequality aversion parameter. As ϵ increases, $A(\epsilon)$ becomes more sensitive to transfers among lower incomes and less sensitive to transfers among top incomes [78]. The limiting form of (22) is $A(1) = 1 - \exp(-MLD)$. (All measures considered here are functions of distributional moments, whose existence depends on conditions assuring the convergence of the appropriate integrals. The Gini coefficient (18) exists if and only if the mean of the distribution $\langle x \rangle = \int_0^\infty xf(x; \alpha, \beta, \kappa) dx$ converges, which is true if and only if $\frac{\alpha}{\kappa} > 1$. According to [89], parametric income distribution models share the existence problem of popular inequality measures).

2.3. Estimation

The κ -generalized distribution’s parameters can be estimated using the maximum likelihood technique, which produces estimators with good statistical properties [90,91]. If sample observations $\mathbf{x} = \{x_1, \dots, x_n\}$ are independent, the likelihood function is as follows:

$$L(\mathbf{x}; \theta) = \prod_{i=1}^n f(x_i; \theta)^{w_i} = \prod_{i=1}^n \left\{ \frac{\alpha}{\beta} \left(\frac{x_i}{\beta} \right)^{\alpha-1} \frac{\exp_\kappa[-(x_i/\beta)^\alpha]}{\sqrt{1 + \kappa^2(x_i/\beta)^{2\alpha}}} \right\}^{w_i}, \quad (23)$$

where $f(x_i; \theta)$ denotes the PDF, $\theta = \{\alpha, \beta, \kappa\}$ the vector of unknown parameters, w_i the weight of the i th observation, and n the sample size. This leads to the problem of solving the partial derivatives with respect to α , β and κ for the log-likelihood function:

$$l(\mathbf{x}; \theta) = \ln[L(\mathbf{x}; \theta)] = \sum_{i=1}^n w_i \ln[f(x_i; \theta)], \quad (24)$$

which is the same as finding the solution to the following nonlinear system of equations:

$$\sum_{i=1}^n w_i \frac{\partial}{\partial \alpha} \ln[f(x_i; \theta)] = 0, \quad (25)$$

$$\sum_{i=1}^n w_i \frac{\partial}{\partial \beta} \ln[f(x_i; \theta)] = 0, \quad (26)$$

$$\sum_{i=1}^n w_i \frac{\partial}{\partial \kappa} \ln[f(x_i; \theta)] = 0. \quad (27)$$

However, the derivation of explicit expressions for maximum likelihood estimators of the three κ -generalized parameters poses a challenge due to the absence of feasible analytical solutions. The utilization of numerical optimization algorithms becomes therefore imperative in order to solve the maximum likelihood estimation problem.

3. Application to the Income Distribution in Greece

To celebrate 20 years of Kaniadakis' contribution, it seems appropriate to consider the income distribution in his native Greece to demonstrate the κ -generalized model's capacity to fit real-world data. First, income data for parameter estimation are briefly described. Next, the κ -generalized distribution is fitted to Greek household income data. Finally, using the same income microdata, different income size distribution models are compared.

3.1. Description of the Income Data

Income distribution data for Greece were obtained from the Luxembourg Income Study (LIS) database, which provides public access to household-level data files for various countries, including both developed and developing economies. The data are remote-accessible, requiring program code to be sent to LIS rather than being run directly by the user. At the time of writing, LIS contains Greek income distribution data for the following years: 1995, 2000, 2004, 2007, 2010, 2013, and 2016. The data set used for this review is the 2016 data set based on the 2017 wave of the Greek EU-SILC survey conducted by the Hellenic Statistical Authority (ELSTAT). (EU-SILC is a cross-sectional and longitudinal sample survey coordinated by Eurostat, focusing on income, poverty, social exclusion, and living conditions in the European Union.) The sample size is 22,555 households.

The definition of income is "household disposable income", which is the income available to households to support consumption expenditure and saving during the reference period. The measure includes income from work, wealth, and direct government benefits, but subtracts direct taxes paid. It does not include sales taxes or noncash benefits, such as healthcare provided by a government or employer. Additionally, the income definition excludes income from capital gains, a significant source of nonwage income for wealthy individuals. As a result, many top incomes are likely to be underestimated.

Household disposable income is expressed in euro and "equivalized", i.e., divided by the square root of household size to adjust for differences in household demographics. Prior to equivalization, top and bottom coding is applied to set limits for extreme values. We also exclude all households with missing disposable income and use person-adjusted weights (the product of the household weights and the number of household members) when generating income indicators for the total population and estimating model parameters.

3.2. Results of Fitting

Figure 3 shows the results of fitting the κ -generalized distribution to empirical income data corresponding to the distribution of household income in Greece for the year 2016.

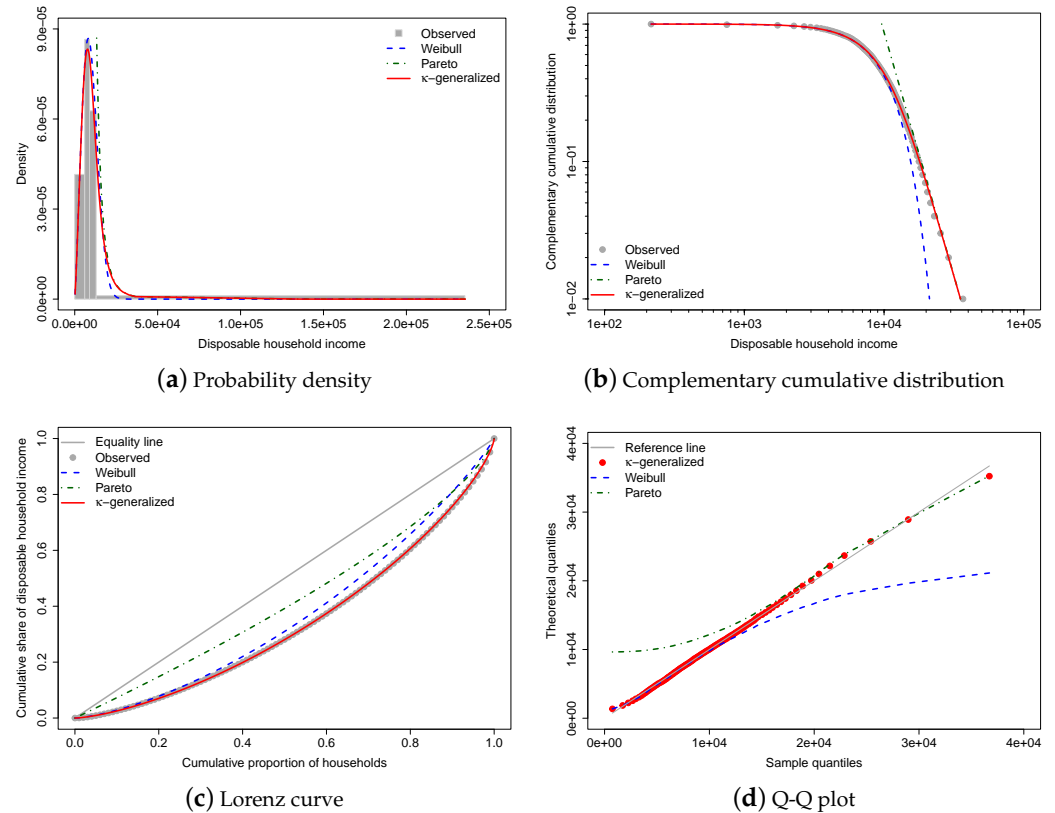


Figure 3. κ -generalized distribution fitted to Greek household income data for 2016. The red solid line represents the κ -generalized model, which fits the data well over the whole range from low to high incomes, including the middle income region. It is compared to the Weibull (blue dashed line) and Pareto power-law (green dashed line) distributions. The complementary cumulative distribution is plotted on double-log axes, emphasizing the right-tail behavior of the distribution. The Lorenz curve plot compares the empirical and theoretical curves, with the gray solid line representing the Lorenz curve of a society with equal income distribution. The Q-Q plot of sample percentiles versus theoretical percentiles of the fitted κ -generalized shows excellent fit, with corresponding percentiles being close to the 45° line from the origin.

The best-fitting parameter values were determined using maximum likelihood estimation, resulting in estimates of $\alpha = 2.233 \pm 0.017$, $\beta = 10,667 \pm 46$, and $\kappa = 0.630 \pm 0.014$. The small errors indicate accurate estimations, and the comparison between the observed and fitted probabilities in panels (a) and (b) of Figure 3 suggests that the κ -generalized distribution has great potential for describing data across the range of low-to-middle-income to high-income power-law regimes, including the intermediate region where Weibull and Pareto distributions show clear departures. (In Figure 3, the curves for the Pareto and Weibull distributions have been drawn by expressing their parameters in terms of the estimated κ -generalized parameters—see Section 2.1).

Panel (c) of the same figure displays data points for the empirical Lorenz curve superimposed on the theoretical curve given by Equation (16) with estimates replacing α and κ as necessary. This formula, represented by the red solid line in the plot, matches the data exceptionally well. In addition, the plot contrasts the empirical Lorenz curve with

the theoretical curves associated with the Weibull and Pareto distributions, respectively, given by:

$$\lim_{\kappa \rightarrow 0} L(u) = P\left(1 + \frac{1}{\alpha}, -\ln(1 - u)\right), \tag{28}$$

where $P(\cdot, \cdot)$ is the lower regularized incomplete gamma function, and:

$$\lim_{x \rightarrow \infty} L(u) = 1 - (1 - u)^{1 - \frac{1}{a}}. \tag{29}$$

As one can easily see, these curves tell only a small part of the story.

To provide an indirect check on the validity of the parameter estimation, we have also computed predicted values for median and mean household disposable income, as well as the Gini and Atkinson coefficients—the latter with the inequality aversion parameter ϵ equal to 1. The results, obtained by substituting the estimated parameters into relevant expressions, are presented in Table 1, along with their empirical counterparts, corresponding to the LIS staff’s “Inequality and Poverty Key Figures” for the considered country and year. (In this article, inequality measures are calculated using the most recent version of DASP, the Distributive Analysis Stata Package [92], which is available at <http://dasp.ecn.ulaval.ca/>—accessed on 29 June 2023. The complete set of corresponding “key figures” is available in an Excel workbook that can be downloaded from <https://www.lisdatacenter.org/data-access/key-figures/>—accessed on 29 June 2023.)

Table 1. Observed and predicted values of the median, the mean, the Gini index G and the Atkinson inequality measure $A(1)$.

Statistic	Observed			Predicted
	Value	LB ^a	UB ^b	
Median	9123	8983	9264	9181
Mean	10,548	10,292	10,805	10,488
G	0.323	0.312	0.334	0.322
$A(1)$	0.179	0.169	0.189	0.172

Notes: ^a lower bound of the 95% normal-based confidence interval obtained by adding -1.96 times the standard error to the sample indicator; ^b upper bound of the 95% normal-based confidence interval obtained by adding $+1.96$ times the standard error to the sample indicator. Source: author’s calculations based on Greek LIS data for 2016.

The κ -generalized distribution predictions are fully covered by asymptotic normal 95% confidence intervals, confirming excellent agreement between the model and sample observations.

The linear behavior of the quantile-quantile (Q-Q) plot of sample percentiles against the fitted κ -generalized distribution and its limiting cases, shown in panel (d) of Figure 3, confirms the model’s validity as well as the fact that the Weibull and Pareto distributions provide partial and incomplete data descriptions.

3.3. Comparisons of Alternative Distributions

This section compares the κ -generalized distribution’s performance with other parametric models, including the three-parameter generalized gamma [93], Singh–Maddala [33], and Dagum type I [34] distributions, which have the following PDFs, respectively:

$$f(x; a, \beta, p) = \frac{ax^{ap-1} \exp[-(x/\beta)^a]}{\beta^{ap}\Gamma(a)}, \quad x > 0, \quad a, \beta, p > 0, \tag{30}$$

$$f(x; a, b, q) = \frac{aqx^{a-1}}{b^a [1 + (x/b)^a]^{1+q}}, \quad x > 0, \quad a, b, q > 0, \tag{31}$$

$$f(x; a, b, p) = \frac{apx^{ap-1}}{b^{ap}[1 + (x/b)^a]^{p+1}}, \quad x > 0, \quad a, b, p > 0. \tag{32}$$

Ref. [77] provides analytical expressions for distribution functions, moments, and tools for inequality measurement, including the Lorenz curve and Gini coefficient. Refs. [87,94] provide formulas for generalized entropy measures of the GB2 distribution, from which the Singh–Maddala and Dagum versions are easily obtained. For the generalized gamma distribution, closed expressions for the Theil entropy index and the mean logarithmic deviation are given in Refs. [85,95]. (Let X be a random variable following the generalized beta distribution of the second kind (GB2) with parameters a, b, p , and q , i.e., $X \sim \text{GB2}(a, b, p, q)$). The Singh–Maddala distribution is the special case of the GB2 distribution when $p = 1$; the Dagum type I distribution is the special case when $q = 1$. For a discussion of other special cases, see [35,77]).

Table 2 displays maximum likelihood estimates for the models under consideration.

Table 2. Maximum likelihood estimates for the generalized gamma, Singh–Maddala, Dagum type I and κ -generalized models of income distribution.

Model ^a	Parameters ^b			Goodness-of-Fit Criteria ^{c,d}			
	a (α)	b (β)	q, p, κ	RMSE	MAE	LRMSE	LMAE
GG	0.684 (0.018)	829 (115)	5.475 (0.270)	2.325	2.047	0.939	0.812
SM	2.441 (0.021)	12,531 (219)	1.835 (0.053)	0.716	0.574	0.319	0.187
D	3.705 (0.041)	11,705 (104)	0.560 (0.011)	0.539	0.437	0.281	0.211
κ -gen	2.233 (0.017)	10,667 (46)	0.630 (0.014)	0.530	0.418	0.188	0.139

Notes: ^a GG = generalized gamma, SM = Singh-Maddala, D = Dagum type I, κ -gen = κ -generalized; ^b numbers in parentheses: estimated standard errors; ^c RMSE = root mean square error, MAE = mean absolute error, LRMSE = root mean square error between the observed and estimated Lorenz curves, LMAE = mean absolute error between the observed and estimated Lorenz curves; ^d values multiplied by 100. Source: author’s calculations based on Greek LIS data for 2016.

The κ -generalized model offers the best results, with parameter standard errors derived from the inverse Hessian matrix being the lowest among competing income distribution models.

The root mean square error and mean absolute error between observed and predicted probabilities were used to determine which distribution best fits the data. These goodness-of-fit measures are, respectively, defined by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [\hat{F}_W(t) - F(x_i; \hat{\theta})]^2} \tag{33}$$

and:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{F}_W(t) - F(x_i; \hat{\theta})|, \tag{34}$$

where $\hat{F}_W(t) = \frac{1}{W} \sum_{i=1}^n w_i \mathbb{1}\{x_i \leq t\}$, with $W = \sum_{i=1}^n w_i$, denotes the weighted empirical cumulative distribution function—equal to the sum of the income weights where $x \leq t$ divided by the total sum of weights—and $\hat{\theta}$ is the vector of estimated parameters. (In the formulas above, $\mathbb{1}\{\cdot\}$ is an indicator function that takes the value 1 if the condition in $\{\cdot\}$ is true, 0 otherwise.) The RMSE and MAE between the observed and estimated Lorenz

curves have also been used as goodness-of-fit criteria, as they are expected to better reflect the accuracy of the inequality estimates. These additional measures are given by:

$$LRMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [L_i - L(\lambda_i; \hat{\theta})]^2} \tag{35}$$

and:

$$LMAE = \frac{1}{n} \sum_{i=1}^n |L_i - L(\lambda_i; \hat{\theta})|, \tag{36}$$

where $\lambda_i = \hat{F}_W(t)$ and L_i denote the cumulative share of population and income, respectively, up to percentile i —i.e., (λ_i, L_i) is a point on the empirical Lorenz curve.

Based on the above goodness-of-fit criteria, the κ -generalized model is clearly the best fit. As shown in the last three columns of Table 2, the generalized gamma, Singh–Maddala, and Dagum type I have larger *RMSE* and *MAE* values for both probabilities and Lorenz curves, suggesting that these models perform worse than the κ -generalized distribution.

The performance of the four models is further evaluated by considering the accuracy of selected distributional statistics implied by parameter estimates. Table 3 presents the predicted values for the median, mean, and several inequality measures derived from estimates in Table 2. (The Gini coefficient of the generalized gamma distribution is available in [35] as a long expression involving the Gaussian hypergeometric function ${}_2F_1$, which is not currently available in the online statistical evaluator provided by the LIS web-based interface. An estimate of the Gini index for the generalized gamma distribution was therefore obtained by numerically integrating the area between the predicted Lorenz curve and the line of hypothetical equality. Ref. [96] reviews various methods for numerically estimating the Gini.)

Table 3. Observed and predicted values of selected distributional statistics.

Statistic ^a	Observed			Predicted ^d			
	Value	LB ^b	UB ^c	GG	SM	D	κ -gen
Median	9123	8983	9264	9076	9108	9189	9181
Mean	10,548	10,292	10,805	10,532	10,463	10,447	10,488
<i>G</i>	0.323	0.312	0.334	0.333	0.321	0.320	0.322
<i>MLD</i>	0.197	0.185	0.209	0.196	0.183	0.188	0.188
<i>T</i>	0.191	0.175	0.207	0.180	0.174	0.178	0.181
<i>A</i> (1)	0.179	0.169	0.189	0.178	0.168	0.172	0.172

Notes: ^a *G* = Gini index, *MLD* = mean logarithmic deviation index, *T* = Theil index, *A*(1) = Atkinson coefficient with inequality aversion parameter ϵ equal to 1; ^b lower bound of the 95% normal-based confidence interval obtained by adding -1.96 times the standard error to the sample indicator; ^c upper bound of the 95% normal-based confidence interval obtained by adding $+1.96$ times the standard error to the sample indicator; ^d GG = generalized gamma, SM = Singh–Maddala, D = Dagum type I, κ -gen = κ -generalized. Source: author’s calculations based on Greek LIS data for 2016.

For each of the models examined, the accuracy of the implied statistics is evaluated by calculating the absolute percentage error:

$$APE = \frac{|P - A|}{A} \times 100 \tag{37}$$

between the predicted values (*P*) and the actual sample estimates (*A*) given in Table 3. The results are summarized in Figure 4.

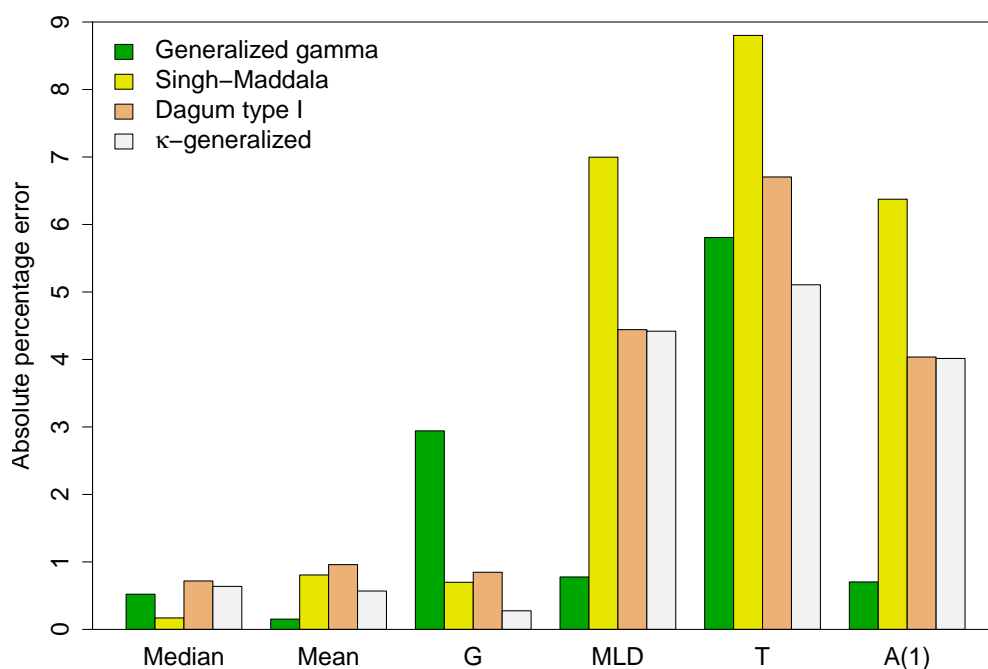


Figure 4. Absolute percentage error between the predicted values for key distributional summary measures and their sample counterparts.

Except for the median, the κ -generalized distribution has more accurate implied estimates of selected distributional statistics than the Singh–Maddala and Dagum type I models, with the Gini coefficient being significantly more accurate. This implies that the κ -generalized estimation procedure preserves the mean characteristic of the analyzed data and accurately models intra- and/or inter-group variation. Additionally, when considering income differences in different parts of the income distribution, the κ -generalized provides more accurate estimates than the two competitors of the *MLD* index, Theil index *T* and the Atkinson inequality measure *A*(1). The Gini is an inequality index sensitive to the middle, while the other indices are more sensitive to the top and bottom of the income distribution. These results support the closest approximation to the income distribution found for the κ -generalized model.

The κ generalized distribution also outperforms the generalized gamma in predicting the Gini coefficient and Theil index, while the generalized gamma provides more accurate estimates for the *MLD* index, the *A*(1) measure, the median, and the mean. This agreement is due to better fit in the lower part of the observed distribution, while disagreements arise from poorer fit in the upper-middle range, especially at the top end. This is demonstrated by the double-logarithmic plot in Figure 5, known as the Zipf plot, which shows the relationship between income and the complementary CDF of income for the data under study.

The Zipf plot is natural to use when looking at the upper part of the distribution because it puts more emphasis on the upper tail and makes it easier to detect deviations in that part of the distribution from what a model would predict [97]. The lines show the Zipf plots that were predicted by fitting the generalized gamma and κ -generalized models. As the graph shows, both are pretty close to the actual data in the lower part of the income distribution. However, the empirical observations of the upper tail are very different from what the generalized gamma says they should be, while the theoretical Zipf plot for the κ -generalized distribution is much closer to the empirical one in the same part of the observed income distribution.

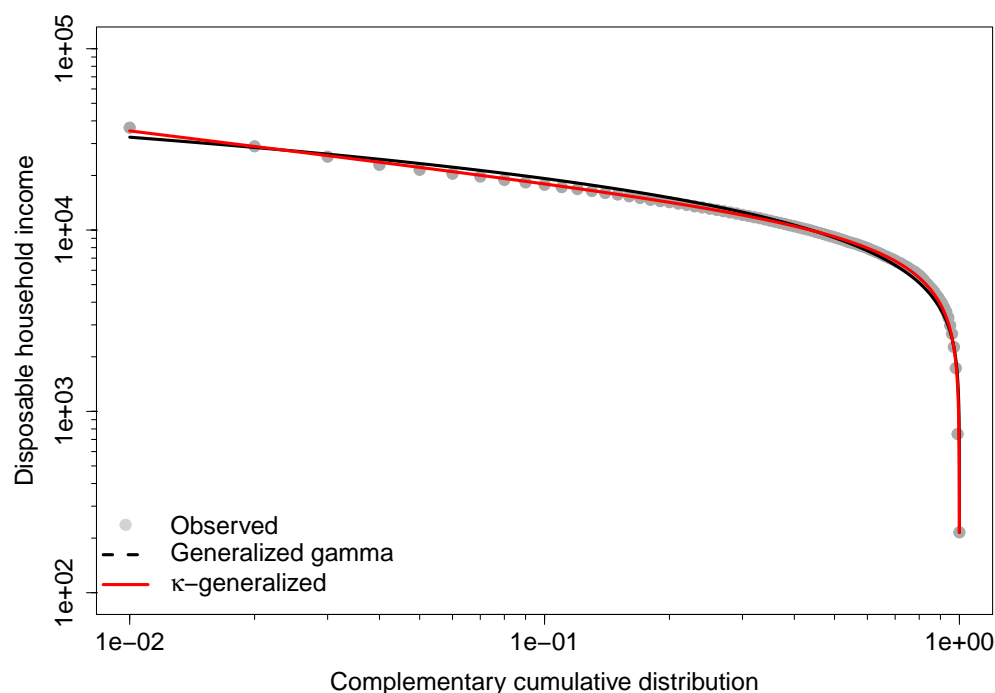


Figure 5. Zipf plot for the 2016 Greek household income data. The lines are the predicted Zipf plots obtained from the fit of the generalized gamma and κ -generalized models.

4. Applications of κ -Generalized Models to Income and Wealth Data

Apart from the one considered in this review, there have been numerous applications of the κ -generalized model to real-world income data over the past two decades.

The first study was conducted by [47], who analyzed 2001–2002 household incomes in Germany, Italy, and the United Kingdom. They found excellent agreement between the model and the empirical distributions across the full spectrum of incomes, including the intermediate income range where clear deviation was found when the Weibull model and pure Pareto law were used for interpolation.

The κ -generalized distribution was later applied to Australian household incomes in 2002–2003 [56] and US family incomes in 2003 [56,57]. The model again described the entire income range well and accurately estimated the inequality level in both countries using the Lorenz curve and Gini measure.

Comparative studies that fit multiple distributions to the same data are crucial for comparing performance. For example, Ref. [58], which examined the distribution of household income in Italy from 1989 to 2006, showed that the κ -generalized model outperforms three-parameter competitors such as the Singh–Maddala and Dagum type I distributions, except for the GB2, which has an extra parameter. The model has also been used to analyze household income data for Germany between 1984 and 2007, the United Kingdom between 1991 and 2004, and the United States between 1980 and 2005. In many cases, the distribution of household income is observed to conform to the κ -generalized model, rather than the Singh–Maddala or Dagum type I distributions. In particular, the κ -generalized distribution is found to outperform competitors in the right tail of the data. The three-parameter κ -generalized model provides superior income inequality estimates even when the fit is worse than distributions belonging to the GB2 family, as obtained by [98] when comparing US and Italian income data for the 2000s. Finally, Ref. [60] finds that the κ -generalized distribution offers a superior fit to the data and, in many cases, estimates income inequality more accurately than alternatives using household income data for 45 countries from Wave IV to Wave IX of the LIS database. (Four-parameter extensions of the κ -generalized distribution, called *extended κ -generalized distributions of the first and second kind*— $E\kappa G1$ and $E\kappa G2$, respectively—were introduced by [74]. These two

extensions are not discussed here, but Refs. [60,61,74] provide formulas for the moments, Lorenz curve, Gini index, coefficient of variation, mean logarithmic deviation, and Theil index for both the models. The new variants of the κ -generalized distribution outperform other four-parameter models in almost all cases, especially in estimating inequality indices with greater precision. In addition, a κ -deformation of the generalized gamma distribution with a power-law tail has recently been proposed by [99], to which the reader is referred for further details.)

The κ -generalized distribution has also been used to analyze the singularities of survey data on *net* wealth, which is gross wealth minus total debt [60,61,100]. These data show highly significant frequencies of households or individuals with wealth that is either null or negative. The κ -generalized model of wealth distribution is a mixture of an atomic and two continuous distributions. The atomic distribution accounts for economic units with no net worth, while a Weibull function accounts for negative net worth data. Positive net worth values, on the other hand, are represented by the κ -generalized model (3). The κ -generalized mixture model for wealth distribution was used to model US net worth data from 1984 to 2011 [100]. The model was generally accurate and its performance was superior to that of finite mixture models based on the Singh–Maddala and Dagum type I distributions for positive net worth values. Similar results were later obtained by Ref. [60] when analyzing net wealth data for nine countries selected from the Luxembourg Wealth Study (LWS) database. (The Luxembourg Wealth Study database—see <https://www.lisdatacenter.org/our-data/lws-database/>, accessed on 29 June 2023—is a collaborative project to assemble existing microdata on household wealth into a coherent database, aiming to do for wealth what the LIS database has achieved for income. The LWS was officially launched in 2004 and currently provides wealth data sets for several countries and years.)

5. Concluding Remarks

The κ -generalized distribution, a statistical model developed over several years of collaborative, multidisciplinary research, is a valuable tool for studying income and wealth distributions. This article discussed its basic properties, relationships with other distributions, and important extensions. It also discussed common inequality measures such as the Lorenz curve and Gini index, and how they can be computed from κ -generalized parameter estimates. A review of empirical applications showed excellent agreement with observed data. It is hoped that the collection of all these results in a single source will facilitate and promote the use of the κ -generalized distribution.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The Luxembourg Income Study Database (<https://www.lisdatacenter.org/>, accessed on 26 June 2023) provides remote access to the microdata through a web-based Job Submission Interface (LISSY). Users have to register to the platform and submit through the LISSY interface their statistical programs written in R, SAS, SPSS or Stata. Data analysis was performed using Stata software version 17 [101] while graphs were generated using R software version 4.3.1 [102]. To allow reproduction of the analysis, software code used in this article is available from the author on request.

Acknowledgments: The author would like to thank the anonymous referees whose comments and suggestions made it possible to improve this paper greatly.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Chatterjee, A.; Yarlagadda, S.; Chakrabarti, B.K. *Econophysics of Wealth Distributions*; Springer: Milano, Italia, 2005.
2. Yakovenko, V.M. Econophysics, statistical mechanics approach to. In *Encyclopedia of Complexity and Systems Science*; Meyers, R.A., Ed.; Springer: New York, NY, USA, 2009; pp. 2800–2826.
3. Yakovenko, V.M.; Rosser, J.B. *Colloquium: Statistical mechanics money, wealth, income*. *Rev. Mod. Phys.* **2009**, *81*, 1703–1725. [[CrossRef](#)]
4. Chakrabarti, B.K.; Chakraborti, A.; Chakravarty, S.R.; Chatterjee, A. *Econophysics of Income and Wealth Distributions*; Cambridge University Press: New York, NY, USA, 2013.
5. Milanovic, B. *The Haves and the Have-Nots: A Brief and Idiosyncratic History of Global Inequality*; Basic Books: New York, NY, USA, 2011.
6. Stiglitz, J.E. *The Price of Inequality: How Today's Divided Society Endangers Our Future*; W. W. Norton & Company: New York, NY, USA, 2012.
7. Piketty, T. *Capital in the Twenty-First Century*; The Belknap Press of Harvard University Press: Cambridge, MA, USA, 2014.
8. Atkinson, A.B. *Inequality: What Can Be Done?* Harvard University Press: Cambridge, MA, USA, 2015.
9. Stiglitz, J.E. *The Great Divide: Unequal Societies and What We Can Do about Them*; W. W. Norton & Company: New York, NY, USA, 2015.
10. Pareto, V. La legge della domanda. *G. Degli Econ.* **1895**, *10*, 59–68.
11. Pareto, V. La courbe de la répartition de la richesse. In *Recueil Publié par la Faculté de Droit à l'Occasion de l'Exposition Nationale Suisse*; Viret-Genton, C., Ed.; Université de Lausanne: Lausanne, Switzerland, 1896; pp. 373–387.
12. Pareto, V. *Cours d'Économie Politique*; Macmillan: London, UK, 1897.
13. Pareto, V. Aggiunta allo studio della curva delle entrate. *G. Degli Econ.* **1897**, *14*, 15–26.
14. Montroll, E.W.; Shlesinger, M.F. On $1/f$ noise and other distributions with long tails. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 3380–3383. [[CrossRef](#)] [[PubMed](#)]
15. Montroll, E.W.; Shlesinger, M.F. Maximum entropy formalism, fractals, scaling phenomena, and $1/f$ noise: A tale of tails. *J. Stat. Phys.* **1983**, *32*, 209–230. [[CrossRef](#)]
16. Drăgulescu, A.; Yakovenko, V.M. Evidence for the exponential distribution of income in the USA. *Eur. Phys. J. B* **2001**, *20*, 585–589. [[CrossRef](#)]
17. Drăgulescu, A.; Yakovenko, V.M. Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States. *Phys. A Stat. Mech. Its Appl.* **2001**, *299*, 213–221. [[CrossRef](#)]
18. Souma, W. Universal structure of the personal income distribution. *Fractals* **2001**, *9*, 463–470. [[CrossRef](#)]
19. Clementi, F.; Gallegati, M. Power law tails in the Italian personal income distribution. *Phys. A Stat. Mech. Its Appl.* **2005**, *350*, 427–438. [[CrossRef](#)]
20. Clementi, F.; Gallegati, M. Pareto's law of income distribution: Evidence for Germany, the United Kingdom, and the United States. In *Econophysics of Wealth Distributions*; Chatterjee, A., Yarlagadda, S., Chakrabarti, B.K., Eds.; Springer: Milan, Italy, 2005; pp. 3–14.
21. Silva, A.C.; Yakovenko, V.M. Temporal evolution of the “thermal” and “superthermal” income classes in the USA during 1983–2001. *Europhys. Lett.* **2005**, *69*, 304–310. [[CrossRef](#)]
22. Nirei, M.; Souma, W. A two factor model of income distribution dynamics. *Rev. Income Wealth* **2007**, *53*, 440–459. [[CrossRef](#)]
23. Gibrat, R. *Les Inégalités économiques. Applications: Aux Inégalités des Richesses, à la Concentration des Entreprises, aux Populations des Villes, aux Statistiques des Familles, etc., d'une loi Nouvelle: La loi de l'Effet Proportionnel*; Librairie du Recueil Sirey: Paris, France, 1931.
24. Salem, A.B.Z.; Mount, T.D. A convenient descriptive model of income distribution: The gamma density. *Econometrica* **1974**, *42*, 1115–1127. [[CrossRef](#)]
25. Aitchison, J.; Brown, J.A.C. On criteria for descriptions of income distribution. *Metroeconomica* **1954**, *6*, 88–107. [[CrossRef](#)]
26. Aitchison, J.; Brown, J.A.C. *The Lognormal Distribution with Special Reference to Its Use in Economics*; Cambridge University Press: New York, NY, USA, 1957.
27. McDonald, J.B.; Ransom, M.R. Functional forms, estimation techniques and the distribution of income. *Econometrica* **1979**, *47*, 1513–1525. [[CrossRef](#)]
28. Majumder, A.; Chakravarty, S.R. Distribution of personal income: Development of a new model and its application to U. S. income data. *J. Appl. Econom.* **1990**, *5*, 189–196. [[CrossRef](#)]
29. Atoda, N.; Suruga, T.; Tachibanaki, T. Statistical inference of functional forms for income distribution. *Econ. Stud. Q.* **1988**, *39*, 14–40.
30. Esteban, J.M. Income-share elasticity and the size distribution of income. *Int. Econ. Rev.* **1986**, *27*, 439–444. [[CrossRef](#)]
31. Kloek, T.; van Dijk, H.K. Efficient estimation of income distribution parameters. *J. Econom.* **1978**, *8*, 61–74. [[CrossRef](#)]
32. Taillie, C. Lorenz ordering within the generalized gamma family of income distributions. In *Statistical Distributions in Scientific Work*; Taillie, C., Patil, G.P., Baldessari, B.A., Eds.; D. Reidel Publishing Company: Dordrecht, The Netherlands, 1981; Volume 6, pp. 181–192.
33. Singh, S.K.; Maddala, G.S. A function for size distribution of incomes. *Econometrica* **1976**, *44*, 963–970. [[CrossRef](#)]
34. Dagum, C. A new model of personal income distribution: Specification and estimation. *Econ. Appliquée* **1977**, *30*, 413–436.
35. McDonald, J.B. Some generalized functions for the size distribution of income. *Econometrica* **1984**, *52*, 647–665. [[CrossRef](#)]

36. Reed, W.J.; Jorgensen, M. The double Pareto-lognormal distribution—A new parametric model for size distributions. *Commun. Stat.-Theory Methods* **2004**, *33*, 1733–1753. [[CrossRef](#)]
37. McDonald, J.B.; Xu, Y.J. A generalization of the beta distribution with applications. *J. Econom.* **1995**, *66*, 133–152. [[CrossRef](#)]
38. Reed, W.J. Brownian-Laplace motion and its use in financial modelling. *Commun. Stat. Theory Methods* **2007**, *36*, 473–484. [[CrossRef](#)]
39. McDonald, J.B.; Ransom, M.R. The generalized beta distribution as a model for the distribution of income: Estimation of related measures of Inequality. In *Modeling Income Distributions and Lorenz Curves*; Chotikapanich, D., Ed.; Springer: New York, NY, USA, 2008; pp. 147–166.
40. Reed, W.J.; Wu, F. New four- and five-parameter models for income distributions. In *Modeling Income Distributions and Lorenz Curves*; Chotikapanich, D., Ed.; Springer: New York, NY, USA, 2008; pp. 211–223.
41. Kaniadakis, G. Non-linear kinetics underlying generalized statistics. *Phys. A Stat. Mech. Its Appl.* **2001**, *296*, 405–425. [[CrossRef](#)]
42. Kaniadakis, G. Statistical mechanics in the context of special relativity. *Phys. Rev. E* **2002**, *66*, 056125. [[CrossRef](#)] [[PubMed](#)]
43. Kaniadakis, G. Statistical mechanics in the context of special relativity. II. *Phys. Rev. E* **2005**, *72*, 036108. [[CrossRef](#)] [[PubMed](#)]
44. Kaniadakis, G. Maximum entropy principle and power-law tailed distributions. *Eur. Phys. J. B* **2009**, *70*, 3–13. [[CrossRef](#)]
45. Kaniadakis, G. Relativistic entropy and related Boltzmann kinetics. *Eur. Phys. J. A* **2009**, *40*, 275–287. [[CrossRef](#)]
46. Kaniadakis, G. Theoretical foundations and mathematical formalism of the power-law tailed statistical distributions. *Entropy* **2013**, *15*, 3983–4010. [[CrossRef](#)]
47. Clementi, F.; Gallegati, M.; Kaniadakis, G. κ -generalized statistics in personal income distribution. *Eur. Phys. J. B* **2007**, *57*, 187–193. [[CrossRef](#)]
48. Rajaonarison, D.; Bolduc, D.; Jayet, H. The K -deformed multinomial logit model. *Econ. Lett.* **2005**, *86*, 13–20. [[CrossRef](#)]
49. Rajaonarison, D. Deterministic heterogeneity in tastes and product differentiation in the K -logit model. *Econ. Lett.* **2008**, *100*, 396–398. [[CrossRef](#)]
50. Trivellato, B. Replication and shortfall risk in a binomial model with transaction costs. *Math. Methods Oper. Res.* **2009**, *69*, 1–26. [[CrossRef](#)]
51. Imparato, D.; Trivellato, B. Geometry of extended exponential models. In *Algebraic and Geometric Methods in Statistics*; Gibilisco, P., Riccomagno, E., Pistone, G., Wynn, H.P., Eds.; Cambridge University Press: Cambridge, UK, 2009; pp. 307–326.
52. Trivellato, B. The minimal κ -entropy martingale measure. *Int. J. Theor. Appl. Financ.* **2012**, *15*, 1250038. [[CrossRef](#)]
53. Trivellato, B. Deformed exponentials and applications to finance. *Entropy* **2013**, *15*, 3471–3489. [[CrossRef](#)]
54. Moretto, E.; Pasquali, S.; Trivellato, B. Option pricing under deformed Gaussian distributions. *Phys. A Stat. Mech. Appl.* **2016**, *446*, 246–263. [[CrossRef](#)]
55. Moretto, E.; Pasquali, S.; Trivellato, B. A non-Gaussian option pricing model based on Kaniadakis exponential deformation. *Eur. Phys. J. B* **2017**, *90*, 179. [[CrossRef](#)]
56. Clementi, F.; Di Matteo, T.; Gallegati, M.; Kaniadakis, G. The κ -generalized distribution: A new descriptive model for the size distribution of incomes. *Phys. A Stat. Mech. Appl.* **2008**, *387*, 3201–3208. [[CrossRef](#)]
57. Clementi, F.; Gallegati, M.; Kaniadakis, G. A κ -generalized statistical mechanics approach to income analysis. *J. Stat. Mech. Theory Exp.* **2009**, *2009*, P02037. [[CrossRef](#)]
58. Clementi, F.; Gallegati, M.; Kaniadakis, G. A model of personal income distribution with application to Italian data. *Empir. Econ.* **2010**, *39*, 559–591. [[CrossRef](#)]
59. Clementi, F.; Gallegati, M.; Kaniadakis, G. A new model of income distribution: The κ -generalized distribution. *J. Econ.* **2012**, *105*, 63–91. [[CrossRef](#)]
60. Clementi, F.; Gallegati, M. *The Distribution of Income and Wealth: Parametric Modeling with the κ -Generalized Family*; Springer International Publishing AG: Cham, Switzerland, 2016.
61. Clementi, F.; Gallegati, M.; Kaniadakis, G.; Landini, S. κ -generalized models of income and wealth distributions: A survey. *Eur. Phys. J. Spec. Top.* **2016**, *225*, 1959–1984. [[CrossRef](#)]
62. Clementi, F.; Gallegati, M. New economic windows on income and wealth: The κ -generalized family of distributions. *J. Soc. Econ. Stat.* **2017**, *6*, 1–15.
63. Landini, S. Mathematics of *strange* quantities: Why are κ -generalized models a good fit to income and wealth distributions? An explanation. In *The Distribution of Income and Wealth: Parametric Modeling with the κ -Generalized Family*; Clementi, F., Gallegati, M., Eds.; Springer International Publishing AG: Cham, Switzerland, 2016; pp. 93–132.
64. Bartels, C.P.A.; van Metelen, H. *Alternative Probability Density Functions of Income: A Comparison of the Longnormal-, Gamma- and Weibull-Distribution with Dutch Data*; Research Memorandum 29; Department of Quantitative Studies, Faculty of Economics, Vrije Universiteit: Amsterdam, The Netherlands, 1975.
65. Bartels, C.P.A. *Economic Aspects of Regional Welfare: Income Distribution and Unemployment*; Martinus Nijhoff: Leiden, The Netherlands, 1977.
66. Espinguet, P.; Terraza, M. Essai d'extrapolation des distributions de salaires français. *Econ. Appliquée* **1983**, *36*, 535–561.
67. Bordley, R.F.; McDonald, J.B.; Mantrala, A. Something new, something old: Parametric models for the size distribution of income. *J. Income Distrib.* **1996**, *6*, 91–103. [[CrossRef](#)]
68. Brachmann, K.; Stich, A.; Trede, M. Evaluating parametric income distribution models. *Allg. Stat. Arch.* **1996**, *80*, 285–298.

69. Tachibanaki, T.; Suruga, T.; Atoda, N. Estimations of income distribution parameters for individual observations by maximum likelihood method. *J. Jpn. Stat. Soc.* **1997**, *27*, 191–203. [[CrossRef](#)]
70. Mandelbrot, B. The Pareto-Lévy law and the distribution of income. *Int. Econ. Rev.* **1960**, *1*, 79–106. [[CrossRef](#)]
71. Kakwani, N. *Income Inequality and Poverty: Methods of Estimation and Policy Applications*; Oxford University Press: New York, NY, USA, 1980.
72. Lorenz, M.O. Methods of measuring the concentration of wealth. *Publ. Am. Stat. Assoc.* **1905**, *9*, 209–219.
73. Gastwirth, J.L. A general definition of the Lorenz curve. *Econometrica* **1971**, *39*, 1037–1039. [[CrossRef](#)]
74. Okamoto, M. *Extension of the κ -Generalized Distribution: New Four-Parameter Models for the Size Distribution of Income and Consumption*; Working Paper 600; LIS Cross-National Data Center: Luxembourg, 2013. Available online: <https://www.lisdatacenter.org/wps/liswps/600.pdf> (accessed on 26 June 2023).
75. Gini, C. Sulla misura della concentrazione e della variabilità dei caratteri. *Atti Del R. Ist. Veneto Di Sci. Lett. Ed Arti* **1914**, *73*, 1201–1248.
76. Arnold, B.C.; Laguna, L. *On Generalized Pareto Distributions with Applications to Income Data*; Iowa State University Press: Ames, IA, USA, 1977.
77. Kleiber, C.; Kotz, S. *Statistical Size Distributions in Economics and Actuarial Sciences*; John Wiley & Sons: New York, NY, USA, 2003.
78. Allison, P.D. Measures of inequality. *Am. Sociol. Rev.* **1978**, *43*, 865–880. [[CrossRef](#)]
79. Cowell, F.A. Generalized entropy and the measurement of distributional change. *Eur. Econ. Rev.* **1980**, *13*, 147–159. [[CrossRef](#)]
80. Cowell, F.A. On the structure of additive inequality measures. *Rev. Econ. Stud.* **1980**, *47*, 521–531. [[CrossRef](#)]
81. Shorrocks, A.F. The class of additively decomposable inequality measures. *Econometrica* **1980**, *48*, 613–625. [[CrossRef](#)]
82. Cowell, F.A.; Kuga, K. Additivity and the entropy concept: An axiomatic approach to inequality measurement. *J. Econ. Theory* **1981**, *25*, 131–143. [[CrossRef](#)]
83. Cowell, F.A.; Kuga, K. Inequality measurement: An axiomatic approach. *Eur. Econ. Rev.* **1981**, *15*, 287–305. [[CrossRef](#)]
84. Theil, H. *Economics and Information Theory*; North-Holland: Amsterdam, The Netherlands, 1967.
85. Sarabia, J.M.; Jordá, V.; Remuzgo, L. The Theil indices in parametric families of income distributions—A short review. *Rev. Income Wealth* **2017**, *63*, 867–880. [[CrossRef](#)]
86. Atkinson, A.B. On the measurement of inequality. *J. Econ. Theory* **1970**, *2*, 244–263. [[CrossRef](#)]
87. Jenkins, S.P. Distributionally-sensitive inequality indices and the GB2 income distribution. *Rev. Income Wealth* **2009**, *55*, 392–398. [[CrossRef](#)]
88. Cowell, F.A. *Measuring Inequality*; Oxford University Press: New York, NY, USA, 2011.
89. Kleiber, C. The existence of population inequality measures. *Econ. Lett.* **1997**, *57*, 39–44. [[CrossRef](#)]
90. Rao, C. *Linear Statistical Inference and Its Applications*; John Wiley & Sons: New York, NY, USA, 1973.
91. Ghosh, J.K. *Higher Order Asymptotics*; Institute of Mathematical Statistics and American Statistical Association: Hayward, CA, USA, 1994.
92. Araar, A.; Duclos, J.Y. *User Manual for Stata Package DASP: Version 3.0*; PEP, World Bank, UNDP and Université Laval, 2022. Available online: http://dasp.ecn.ulaval.ca/dasp3/manual/DASP_MANUAL_V303.pdf (accessed on 26 June 2023).
93. Stacy, E.W. A generalization of the gamma distribution. *Ann. Math. Stat.* **1962**, *33*, 1187–1192. [[CrossRef](#)]
94. Chotikapanich, D.; Griffiths, W.; Hajargasht, G.; Karunaratne, W.; Rao, D. Using the GB2 income distribution. *Econometrics* **2018**, *6*, 21. [[CrossRef](#)]
95. Jordá, V.; Alonso, J.M. New estimates on educational attainment using a continuous approach (1970–2010). *World Dev.* **2017**, *90*, 281–293. [[CrossRef](#)]
96. Fellman, J. Estimation of Gini coefficients using Lorenz curves. *J. Stat. Econom. Methods* **2012**, *1*, 31–38.
97. Takayasu, H. *Fractals in Physical Science*; Manchester University Press: Manchester, UK, 1991.
98. Okamoto, M. Evaluation of the goodness of fit of new statistical size distributions with consideration of accurate income inequality estimation. *Econ. Bull.* **2012**, *32*, 2969–2982.
99. Vallejos, A.; Ormazábal, I.; Borotto, F.A.; Astudillo, H.F. A new κ -deformed parametric model for the size distribution of wealth. *Phys. Stat. Mech. Appl.* **2019**, *514*, 819–829. [[CrossRef](#)]
100. Clementi, F.; Gallegati, M.; Kaniadakis, G. A generalized statistical model for the size distribution of wealth. *J. Stat. Mech. Theory Exp.* **2012**, *2012*, P12006. [[CrossRef](#)]
101. StataCorp. *Stata Statistical Software: Release 17*; StataCorp: College Station, TX, USA, 2021. Available online: <https://www.stata.com/> (accessed on 26 June 2023).
102. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023. Available online: <https://www.R-project.org/> (accessed on 26 June 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.