κ-generalized models of income and wealth distributions

A survey

F. Clementi, M. Gallegati, G. Kaniadakis, and S. Landini

1 Dipartimento di Scienze Politiche, della Comunicazione e delle Relazioni Internazionali, Università degli Studi di Macerata, Piazza San Vincenzo Maria Strambi 1, I–62100 Macerata, Italy
2 Dipartimento di Scienze Economiche e Sociali, Università Politecnica delle Marche, Piazzale Raffaele Martelli 8, I–60121 Ancona, Italy
3 Dipartimento Scienza Applicata e Tecnologia, Politecnico di Torino, Corso Duca degli Abruzzi 24, I–10129 Torino, Italy
4 IRES Piemonte, Istituto di Ricerche Economico Sociali del Piemonte, Via Nizza 18, I–10125 Torino, Italy

Abstract The paper provides a survey of results related to the “κ-generalized distribution”, a statistical model for the size distribution of income and wealth. Topics include, among others, discussion of basic analytical properties, interrelations with other statistical distributions as well as aspects that are of special interest in the income distribution field, such as the Gini index and the Lorenz curve. An extension of the basic model that is most able to accommodate the special features of wealth data is also reviewed. The survey of empirical applications given in this paper shows the κ-generalized models of income and wealth to be in excellent agreement with the observed data in many cases.

1 Introduction

A renewed interest in studying the distribution of income and wealth has emerged over the last decade in both the physics [1–3] and economics [5–8] communities. The focus has been mostly put on empirical analysis of extensive data sets to infer the exact shape of income and wealth distributions and to design theoretical models that can reproduce them.

A natural starting point in this area of inquiry was the observation that the number of persons in a population whose incomes exceed $x$ is often well approximated by $C x^{-\alpha}$, for some real $C$ and positive $\alpha$, as Pareto [9–12] argued over 100 years ago. Since the early studies of Pareto, numerous empirical works have shown that the power-law tail is an ubiquitous feature of income and wealth distributions. However, even 100 years after Pareto observation, the understanding of the shape of income/wealth distribution is still far to be complete and definitive. This reflects the fact that there are two distributions, one for the rich, following the Pareto power law, and one for the vast majority of people, which appears to be governed by a completely different law.

Over the years, research in the field has considered a wide variety of functional forms as possible models for the size distribution of income and wealth [13], some of which aim at providing a unified framework for the description of real-world data—including the heavy tails present in empirical income and wealth distributions. Among these, the “κ-generalized distribution” [14–19] was found to work remarkably well. First proposed in 2007, and further developed over successive years, this model finds its roots in the framework of κ-generalized statistical mechanics [21–25,29] and has a bulk very close to the Weibull distribution, while its upper tail decays following a Pareto power law for high values of income and wealth, hence providing a kind of compromise between the two descriptions.

Given the renewed focus on distributional issues, and motivated by the feeling that fruitful interaction between the two fields of statistical mechanics and economics may suggest a new path for investigating
economic relations, it is the aim of the present paper to present a survey of parametric models belonging to the "κ-generalized family", a set of fitting functions for the size distribution of income and wealth that were typically found as superior alternatives to other widely used parametric models. It will be given knowledge of their basic statistical properties and evidence of their empirical plausibility, hoping to convince readers that this class of distributions deserves attention and study.

The paper is organized as follows. In view of its importance for the proposed statistical models, Section 2 reviews the relativistic mechanism originating the \( \kappa \)-exponential function and discusses some of its basic properties. A survey of the \( \kappa \)-generalized models of income and wealth distribution, including interrelations with other distributions as well as basic statistical properties and inferential aspects, is provided in Section 3. In Section 4 we present a survey of empirical applications that show how \( \kappa \)-generalized models of income and wealth are in excellent agreement with the observed data in many cases. Section 5 aims at providing an explanation about why \( \kappa \)-generalized models provide a better fit to empirical distributions than other widely used parametric models for the distribution of income and wealth. Finally, in Section 6 some concluding remarks are reported.

2 The \( \kappa \)-exponential function and its relativistic origin

In experimental particle physics it is well known that the particle distribution at low energies is described by the Boltzmann-Maxwell exponential distribution, i.e.

\[
f (E) \sim \exp \left( -\beta E \right).
\]  

On the contrary, at high energies the power-law tailed distribution

\[
f (E) \sim E^{-\kappa} \quad E \to +\infty
\]

has been observed systematically in the last six decades. This dichotomy in the behavior of the distribution function at low and high energies has stimulated the theoretical research to find a distribution function able to describe the experimental observations on the entire range of energies. Clearly this distribution function had to be motivated by solid and indisputable first principles. In the early 2000s a new proposal was advanced which immediately attracted the attention of many researchers [22–31]. The new proposed distribution is given by

\[
f (E) = N \exp_{\kappa} (-\beta E + \beta \mu),
\]

where \( \beta = 1/K_B T \) is proportional to the reciprocal of the system temperature, \( \mu \) is the chemical potential, and the \( \kappa \)-exponential function \( \exp_{\kappa} (x) \) is defined as

\[
\exp_{\kappa} (x) = \left( \sqrt{1 + \kappa^2 x^2} + \kappa x \right)^{1/\kappa}, \quad 0 < \kappa^2 < 1.
\]

The new distribution at low energies behaves like Boltzmann-Maxwell distribution (1) while at high energies present the power-law tails (2). The more interesting feature of the new distribution (3) is that the roots of \( \kappa \)-exponential function \( \exp_{\kappa} (x) \) given in (4) can be found easily within the Einstein special relativity. In other words, the function \( \exp_{\kappa} (x) \) emerges as a relativistic deformation of the Euler ordinary exponential \( \exp(x) \) of the linear classical physics. The deformation parameter \( \kappa \) is proportional to the reciprocal of light speed \( c \), so that the mechanism generating the \( \exp_{\kappa} (x) \) is simply the one established by the Lorentz transformations of special relativity.

In the following we briefly recall the relativistic mechanism originating the \( \kappa \)-exponential function and discuss its main mathematical properties [22–25,29]. Let us consider in the one-dimension inertial frame \( S \) two identical particles \( A \) and \( B \) of rest mass \( m \). We suppose that the two particles move, in opposite directions, with velocities \( v_A \) and \( v_B \) respectively. The momenta of the two particles are indicated with \( p_A = p (v_A) \) and \( p_B = p (v_B) \) respectively, where

\[
p (v) = \frac{mv}{\sqrt{1 - v^2/c^2}}.
\]

We consider now the same particles in the rest frame \( S' \) of particle \( B \). In this new frame the velocity and momentum of the particle \( B \) are \( v_B' = 0 \) and \( p_B' = 0 \) respectively. In \( S' \) the velocity of particle \( A \) is given by the Einstein law

\[
v_A' = \frac{v_A + v_B}{1 + v_{AUB}/c^2},
\]
defining the relativistic velocity composition law, which follows directly from the kinematic Lorentz transformations.

In order to obtain the additivity law for the relativistic momenta, we first invert (5) so as to have

\[ v = \frac{p/m}{\sqrt{1 + p^2/m^2c^2}}. \]  

(7)

Then, after substitution of the velocity (7) in the velocity additivity law expressed by (6), we obtain the following additivity law for the relativistic momenta

\[ p'_A = p_A \sqrt{1 + \frac{p_B^2}{m^2c^2}} + p_B \sqrt{1 + \frac{p_A^2}{m^2c^2}}. \]  

(8)

Let us now introduce in place of the dimensional variables \((v, p)\) the dimensionless variables \((u, q)\) through

\[ \frac{v}{u} = \frac{p}{mq} = |\kappa| c = v_s < c, \]  

with \(0 < \kappa^2 < 1\). It is easy to verify that the additivity law (8) in terms of the dimensionless momenta \(q\) assumes the form

\[ q'_A = q_A \sqrt{1 + \kappa^2 q_B^2} + q_B \sqrt{1 + \kappa^2 q_A^2}. \]  

(9)

The latter transformation law can be written in the following factorized form

\[ \exp_\kappa (q'_A) = \exp_\kappa (q_A) \exp_\kappa (q_B), \]  

(10)

where the function \(\exp_\kappa (\cdot)\) is the \(\kappa\)-exponential function given by (4). It follows that the \(\kappa\)-exponential can be viewed as the function permitting to write the additivity law (8) of relativistic momenta in the factorized form (10).

From the definition of \(\exp_\kappa (x)\) it results that

\[
\begin{align*}
\exp_0 (x) &\equiv \lim_{\kappa \to 0} \exp_\kappa (x) = \exp (x), \\
\exp_{-\kappa} (x) &\equiv \exp_\kappa (x). 
\end{align*}
\]

(11)

(12)

Like the ordinary exponential, \(\exp_\kappa (x)\) has the properties

\[
\begin{align*}
\exp_\kappa (x) &\in \mathcal{C}^\infty (\mathbb{R}), \\
\frac{d}{dx} \exp_\kappa (x) &> 0, \\
\exp_\kappa (-\infty) &> 0, \\
\exp_\kappa (0) &> 0, \\
\exp_\kappa (\infty) &> 0, \\
\exp_\kappa (x) \exp_\kappa (-x) &= 1.
\end{align*}
\]

(13)

(14)

(15)

(16)

(17)

(18)

The property (18) emerges as particular case of the more general one

\[
\exp_\kappa (x) \exp_\kappa (y) = \exp_\kappa (x \oplus y),
\]

(19)

where

\[ x \oplus y = x \sqrt{1 + \kappa^2 y^2} + y \sqrt{1 + \kappa^2 x^2}. \]  

(20)

Furthermore, \(\exp_\kappa (x)\) has the property

\[
[\exp_\kappa (x)]^r = \exp_{\kappa/r} (rx),
\]

(21)

with \(r \in \mathbb{R}\), which in the limit \(\kappa \to 0\) reproduces one well known property of the ordinary exponential.

We remark the following convexity property

\[
\frac{d^2}{dx^2} \exp_\kappa (x) > 0, \quad x \in \mathbb{R},
\]

(22)

holding when \(\kappa^2 < 1\).
Undoubtedly, one of the more interesting properties of \( \exp_\kappa (x) \) is its power-law asymptotic behavior

\[
\exp_\kappa (x) \sim |2\kappa x|^{1/|\kappa|}.
\]  

(23)

The Taylor expansion of \( \exp_\kappa (x) \) can be written in the following form

\[
\exp_\kappa (x) = \sum_{n=0}^{\infty} \xi_n (\kappa) \frac{x^n}{n!}, \quad \kappa^2 x^2 < 1,
\]

(24)

where the polynomials \( \xi_n (\kappa) \) are given by

\[
\xi_{2m} (\kappa) = \prod_{j=0}^{m-1} \left[ 1 - (2j)^2 \kappa^2 \right], \quad m > 0,
\]

(25)

\[
\xi_{2m+1} (\kappa) = \prod_{j=0}^{m-1} \left[ 1 - (2j+1)^2 \kappa^2 \right], \quad m > 0,
\]

(26)

and can be generated by the following simple recursive formula

\[
\xi_{n+2} (\kappa) = (1 - n^2 \kappa^2) \xi_n (\kappa), \quad n \geq 0,
\]

(27)

\[
\xi_0 (\kappa) = \xi_1 (\kappa) = 1.
\]

(28)

It is remarkable that the first three terms in the Taylor expansion of \( \exp_\kappa (x) \) are the same of the ordinary exponential

\[
\exp_\kappa (x) = 1 + x + \frac{x^2}{2} + (1 - \kappa^2) \frac{x^3}{3!} + (1 - 4\kappa^2) \frac{x^4}{4!}
\]

\[
+ (1 - \kappa^2) (1 - 9\kappa^2) \frac{x^5}{5!} + (1 - 4\kappa^2) (1 - 16\kappa^2) \frac{x^6}{6!} + \cdots.
\]

(29)

3 Use of the \( \kappa \)-generalized statistics for modeling income and wealth distributions

3.1 The \( \kappa \)-generalized model for income distributions

3.1.1 Definitions

In recent years the \( \kappa \)-deformed exponential (4) was adopted successfully to analyze also non-physical systems, including economic systems. In particular, the \( \kappa \)-deformation has been employed to study differentiated product markets [32,33], in finance [26–28,30,31], and for modeling the size distribution of personal income [14–18]. In the latter application a new parametric model, named \( \kappa \)-generalized distribution after [14], was defined in terms of the following cumulative distribution function (CDF)

\[
F (x; \alpha, \beta, \kappa) = 1 - \exp_\kappa \left[ - \left( \frac{x}{\beta} \right)^\alpha \right], \quad x > 0, \quad \alpha, \beta > 0, \quad \kappa \in [0,1),
\]

(30)

where \( x \) denotes the income variable and \( \{ \alpha, \beta, \kappa \} \) are parameters. The corresponding probability density function (PDF) reads as

\[
f (x; \alpha, \beta, \kappa) = \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1} \exp_\kappa \left[ - \left( \frac{x}{\beta} \right)^\alpha \right] \sqrt{1 + \kappa^2 \left( \frac{x}{\beta} \right)^{2\alpha}}.
\]

(31)

References [14–16] give slightly different expressions for the \( \kappa \)-generalized CDF and PDF; their expressions are simply re-parametrized versions of (30) and (31).
3.1.2 Genesis

As shown by [34], the \( \kappa \)-generalized model of income distribution naturally emerges within the field of \( \kappa \)-deformed analysis.

Let \( X: \Omega \to \mathbb{X} = [0 < x_{\min}, x_{\max} \ll +\infty] \) be a random variable denoting income. Let also \( G(x) \in \mathcal{C}^\infty(\mathbb{R}) \) be a strictly increasing, invertible and odd function so as to have \( G(x) \approx x \) for \( x \to x_{\min} \), and let \( \kappa \in [0, 1) \) be a deformation parameter such that \( g_\kappa(x) = G(\kappa x) \) is a \( \kappa \)-deformed generator. The generative representation of the \( \kappa \)-deformed random variable \( X_\kappa \) is a function from \( \mathbb{X} \) into its deformation \( X_\kappa \), namely

\[
X_\kappa \equiv \frac{1}{\kappa} \arcsinh [g_\kappa(X)] = \frac{1}{\kappa} \ln \left[ \sqrt{1 + g_\kappa^2(X)} + g_\kappa(X) \right].
\]  

(32)

Notice that the above is not a proper transformation but rather a deformation, as it aims at deforming the quantity \( X \) to be regularly treated according to \( \kappa \)-deformed operators. Among these, the generalized \( \kappa \)-deformed exponential follows readily from (32) as [21–23,29]

\[
\exp_\kappa(x) \equiv \left[ \sqrt{1 + g_\kappa^2(x) + g_\kappa(x)} \right]^\frac{1}{\kappa} = \exp \left\{ \frac{1}{\kappa} \arcsinh [g_\kappa(x)] \right\} = \exp (x_\kappa).
\]  

(33)

Let now \( \alpha, \beta > 0 \) be two additional parameters and define a regular transformation as

\[
y = \left( \frac{x}{\beta} \right)^\alpha > 0 \implies d\,y = \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1} \,d\,x.
\]  

(34)

Due to self-duality of the \( \kappa \)-deformed exponential, multiplying by \(-1\) and applying the operator (33) one gets

\[
\exp_\kappa(-y) = \sqrt{1 + \kappa^2 y^2} - \kappa y,
\]  

(35)

which defines both the \( \kappa \)-deformed distribution,

\[
U_\kappa (y) = 1 - \exp_\kappa (-y),
\]  

(36)

and its density,

\[
u_\kappa (y) = \frac{\exp_\kappa (-y)}{\sqrt{1 + \kappa^2 y^2}}.
\]  

(37)

Therefore, substituting (34) into (37) gives

\[
u_\kappa \left( \left( \frac{x}{\beta} \right)^\alpha \right) = \frac{d\,U_\kappa \left( \left( \frac{x}{\beta} \right)^\alpha \right)}{\frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1}} \,d\,x = \frac{\exp_\kappa \left[ - \left( \frac{x}{\beta} \right)^\alpha \right]}{\sqrt{1 + \kappa^2 \left( \frac{x}{\beta} \right)^{2\alpha}}}.
\]  

(38)

in view of which it is possible to derive the \( \kappa \)-generalized PDF as the regular derivative of the \( \kappa \)-deformed exponential distribution, i.e.

\[
f (x; \alpha, \beta, \kappa) = \frac{d\,U_\kappa \left( \left( \frac{x}{\beta} \right)^\alpha \right)}{d\,x} = \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1} \exp_\kappa \left[ - \left( \frac{x}{\beta} \right)^\alpha \right] \frac{\exp_\kappa \left( \frac{x}{\beta} \right)^\alpha}{\sqrt{1 + \kappa^2 \left( \frac{x}{\beta} \right)^{2\alpha}}}.
\]  

(39)

3.1.3 Basic properties

As \( \kappa \to 0 \), the \( \kappa \)-generalized distribution tends to the Weibull model with the CDF

\[
\lim_{\kappa \to 0} F(x; \alpha, \beta, \kappa) = 1 - \exp \left[ - \left( \frac{x}{\beta} \right)^\alpha \right],
\]  

(40a)

and the PDF

\[
\lim_{\kappa \to 0} f(x; \alpha, \beta, \kappa) = \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1} \exp \left[ - \left( \frac{x}{\beta} \right)^\alpha \right].
\]  

(40b)
Consequently, the exponential law is also a special limiting case of the $\kappa$-generalized distribution, since it is a special case of the Weibull with $\alpha = 1$.

For $x \to 0^+$ the $\kappa$-generalized behaves similarly to the Weibull model (40), whereas for large $x$ it approaches a Pareto distribution with scale $\beta (2\kappa)^{-\frac{1}{\alpha}}$ and shape $\frac{\alpha}{\kappa}$, i.e.

$$ F(x; \alpha, \beta, \kappa) \sim 1 - \left( \frac{\beta (2\kappa)^{-\frac{1}{\alpha}}}{x} \right)^\frac{\alpha}{\kappa} \quad (41a) $$

and

$$ f(x; \alpha, \beta, \kappa) \sim \frac{\alpha}{\kappa} \left( \frac{\beta (2\kappa)^{-\frac{1}{\alpha}}}{x^{\frac{\alpha}{\kappa} + 1}} \right)^{\frac{\alpha}{\kappa}} \quad (41b) $$

thus satisfying the weak Pareto law [44].

Figures 1 to 3 illustrate the behavior of the $\kappa$-generalized PDF and complementary CDF, $1 - F(x; \alpha, \beta, \kappa)$, for various parameter values. Each graph keeps two parameters constant and varies the remaining one.

The constant $\beta$ is a characteristic scale having the same dimension of income: if $\beta$ is small, then the distribution will be more concentrated around the mode; if $\beta$ is large, then it will be more spread out (Figure 1). Since it is related to the income measurement unit, increments (reductions) in the monetary unit lead to a corresponding augmentation (diminishment) in the value of $\beta$ and generate a global increase (decrease) of each and every one of the incomes and, therefore, of the average income.

The exponent $\alpha$, instead, quantifies the curvature (shape) of the distribution, which is less (more) pronounced for lower (higher) values of the parameter, as seen in Figure 2. It should be noted that for $\alpha \leq 1$ the density exhibits a pole at the origin, whereas for $\alpha > 1$ there exists an interior mode. In the latter case, this mode is at

$$ x_{\text{mode}} = \beta \left( \frac{\alpha^2 + 2\kappa^2 (\alpha - 1)}{2\kappa^2 (\alpha^2 - \kappa^2)} \right)^\frac{1}{\alpha} \left\{ \frac{1 + \frac{4\kappa^2 (\alpha^2 - \kappa^2) (\alpha - 1)^2}{[\alpha^2 + 2\kappa^2 (\alpha - 1)]^2}}{x_{\text{mode}}^{\frac{\alpha}{\kappa} + 1}} - 1 \right\}^{\frac{1}{\alpha}}. \quad (42) $$

The Weibull distribution has received maximum attention in the engineering literature. In physics, it is known as the stretched exponential distribution if $\alpha < 1$. In the economic literature the Weibull is probably less prominent, but reference [35] noticed its potential for income data—although it has been used only sporadically as an income distribution (some applications can be found in references [36–43]).
Figure 2. Plot of the $\kappa$-generalized PDF (a) and CCDF (b) for some different values of $\alpha$ (= 1.00, 1.50, 2.00, 2.50) and fixed $\beta$ (= 1.20) and $\kappa$ (= 0.75). The CCDF is plotted on doubly-logarithmic axes, which is the standard way of emphasizing the right-tail behavior of a distribution. Notice that the curvature (shape) of the distribution becomes less (more) pronounced when the value of $\alpha$ decreases (increases). The case $\alpha = 1.00$ corresponds to the standard exponential distribution.

Figure 3. Plot of the $\kappa$-generalized PDF (a) and CCDF (b) for some different values of $\kappa$ (= 0.00, 0.25, 0.50, 0.75) and fixed $\alpha$ (= 2.00) and $\beta$ (= 1.20). The CCDF is plotted on doubly-logarithmic axes, which is the standard way of emphasizing the right-tail behavior of a distribution. Notice that the upper tail of the distribution fattens (thins) as the value of $\kappa$ increases (decreases). The case $\kappa = 0.00$ corresponds to the Weibull (stretched exponential) distribution.

This built-in flexibility is an attractive feature in that the model can approximate income distributions, which are usually uni-modal, and wealth distributions, which are zero-modal (see Section 3.2).

Finally, as Figure 3 shows, the deformation parameter $\kappa$ measures the fatness of the upper tail: the larger (smaller) its magnitude, the fatter (thinner) the tail.\footnote{Differently than $\beta$, $\alpha$ and $\kappa$ are scale-free (dimensionless) parameters: changes in the monetary unit yield $\beta$ modifications but leave $\alpha$ and $\kappa$ invariant.}
The $\kappa$-generalized admits a closed-form expression for the quantile function,

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

where $\ln_{\kappa} (\cdot)$ denotes the deformed logarithmic function defined as the inverse function of (4), namely $\ln_{\kappa} [\exp_{\kappa} (x)] = \exp_{\kappa} [\ln_{\kappa} (x)] = x$, and reads as

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

(44)

Hence random numbers from a $\kappa$-generalized distribution can be easily generated via the inversion method.

The $r^{th}$ moment exists for $-\alpha < r < \frac{\alpha}{\kappa}$ and equals

$$E (X^r) = \beta^r (2\kappa)^{-\frac{r}{\kappa}} \frac{\Gamma \left( 1 + \frac{r}{\kappa} \right) \Gamma \left( \frac{1}{\kappa} - \frac{r}{2\kappa} \right)}{\Gamma \left( \frac{1}{\kappa} + \frac{1}{2\kappa} \right)},$$

(45)

where $\Gamma (\cdot)$ denotes the gamma function. Specifically,

$$E (X) = m = \beta (2\kappa)^{-\frac{1}{\kappa}} \frac{\Gamma \left( 1 + \frac{1}{\kappa} \right) \Gamma \left( \frac{1}{\kappa} - \frac{1}{2\kappa} \right)}{\Gamma \left( \frac{1}{\kappa} + \frac{1}{2\kappa} \right)}$$

(46)

is the mean of the distribution and

$$\text{Var} (X) = \beta^2 (2\kappa)^{-\frac{2}{\kappa}} \left\{ \frac{\Gamma \left( 1 + \frac{1}{\kappa} \right) \Gamma \left( \frac{1}{\kappa} - \frac{1}{2\kappa} \right)}{1 + \frac{\kappa}{\alpha}} \left[ \frac{\Gamma \left( 1 + \frac{1}{\kappa} \right) \Gamma \left( \frac{1}{\kappa} - \frac{1}{2\kappa} \right)}{\Gamma \left( \frac{1}{\kappa} + \frac{1}{2\kappa} \right)} \right]^2 \right\}$$

(47)

is the variance.

### 3.1.4 Measuring income inequality using the $\kappa$-generalized distribution

The most widely used tool for analyzing and visualizing income inequality is the Lorenz curve [45], and several indices of income inequality are directly related to this curve, most notably the Gini index [46].

Since the quantile function of the $\kappa$-generalized distribution is available in closed form, its normalized integral, the Lorenz curve

$$L (u) = \frac{1}{m} \int_0^u F^{-1} (t) \, dt, \quad u \in [0, 1],$$

(48)

is also of a comparatively simple form, namely [47]

$$L (u) = I_X \left( 1 + \frac{\kappa}{\alpha} \frac{1}{2\kappa} - \frac{1}{2\alpha} \right), \quad X = 1 - (1 - u)^{2\kappa},$$

(49)

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, $I_X (\cdot, \cdot) = \frac{B_X (\cdot, \cdot)}{B_X (\cdot, \cdot)}$. The curve exists if and only if $\frac{\alpha}{\kappa} > 1$.

For the comparison of estimated income distributions it is often of interest to know the parameter constellations for which Lorenz curves do or do not intersect. A complete analytical characterization for the $\kappa$-generalized distribution was obtained by [17]. Suppose $X_i \sim \kappa$-gen ($\alpha_i, \beta_i, \kappa_i$), $i = 1, 2$. The necessary and sufficient conditions for which the Lorenz curves of $X_1$ and $X_2$ do not intersect and we have $X_1 \leq_L X_2$—i.e. the Lorenz curve of $X_1$ lies nowhere below that of $X_2$ for all $u \in [0, 1]$ and consequently $X_1$ exhibits less inequality than $X_2$ in the Lorenz sense—are

$$X_1 \leq_L X_2 \iff \alpha_1 \geq \alpha_2 \quad \text{and} \quad \frac{\alpha_1}{\kappa_1} \geq \frac{\alpha_2}{\kappa_2},$$

(50)

Figure 4 provides an illustration of (50), showing that the less unequal distribution (in the Lorenz sense) always exhibits lighter tails.
As regards scalar measures of inequality, the most widely used of all such indices, the Gini coefficient, takes the form

\[ G = 1 - \frac{2\alpha + 2\kappa}{2\alpha + \kappa} \frac{\Gamma \left( \frac{1}{\alpha} - \frac{1}{\kappa} \right)}{\Gamma \left( \frac{1}{\alpha} + \frac{1}{\kappa} \right)} \Gamma \left( \frac{1}{\alpha} - \frac{1}{\kappa} \right). \]  

(51)

Using Stirling approximation for the gamma function, \( \Gamma(z) \approx \sqrt{2\pi z} z^{z-\frac{1}{2}} \exp(-z) \), and taking the limit as \( \kappa \to 0 \) in Equation (51), one arrives after some simplification at \( G = 1 - 2^{-\frac{1}{\alpha}} \), which is the explicit form of the Gini coefficient for the Weibull distribution [13]. Since the exponential distribution is a special case of the Weibull with shape parameter equal to 1, one directly determines that for \( \kappa \to 0 \) and \( \alpha = 1 \) the Gini coefficient of exponentially distributed incomes is equal to one half [48].

Although commonly used, the Gini is but one of many measures of inequality that are available, and it incorporates particular assumptions about the way in which income differences in different parts of the distribution are summarized—it tends to be most sensitive to transfers around the middle of the income distribution and least sensitive to transfers among the very rich or the very poor [49]. In other forms of income distribution research, the generalized entropy class of inequality measures [50–54] provides a full range of bottom- to top-sensitive indices that are widely used by analysts to assess inequality in different parts of the income distribution. In terms of the \( \kappa \)-generalized parameters, this class of inequality indices is expressed as

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

(52)

Equation (52) defines a class because the index \( GE(\theta) \) assumes different forms depending on the value assigned to \( \theta \), the parameter that characterizes the sensitivity of \( GE(\theta) \) to income differences in different parts of the distribution—the more positive that \( \theta \) is, the more sensitive is \( GE(\theta) \) to income differences at the top of the distribution; the more negative that \( \theta \) is, the more sensitive is \( GE(\theta) \) to income differences

---

\( ^3 \) As it is clear from Figure 4(a), the region around the origin of the \( \kappa \)-generalized distribution is governed by \( \alpha \), the upper tail by both \( \alpha \) and \( \kappa \). More precisely, increasing \( \kappa \) leads to a thicker upper tail, whereas increasing \( \alpha \) lowers both the tails and yields a greater concentration of probability mass around the peak of the distribution. Hence, the \( \alpha \) and \( \kappa \) modifications affect the shape of the distribution and are related to the inequality of the income distribution—as they enter into several inequality measures that can be derived from the \( \kappa \)-generalized model of income distribution. However, the influence of each parameter on these measures is different. For instance, reference [17] obtained a first interpretation of \( \kappa \) as an "inequality" parameter and \( \alpha \) as an "equality" parameter after verifying that the Gini coefficient (51) increases with the former and decreases with the latter. This interpretation has been further corroborated by [20].

---

Figure 4. Tails (a) and Lorenz curves (b) for two \( \kappa \)-generalized distributions. As shown by the dotted line, the Lorenz curves intersect if the parameters are not chosen according to Equation (50).
at the bottom of the distribution. In applied work, two limiting cases of (52) are of particular interest for inequality measurement: the mean logarithmic deviation index

\[ MLD = \lim_{\theta \to 0} GE(\theta) = \frac{1}{\alpha} \left[ \gamma + \psi\left(\frac{1}{2\alpha}\right) + \ln{(2\kappa)} - \alpha \ln\left(\frac{\beta}{m}\right) + \kappa \right], \quad (53) \]

where \( \gamma = -\psi(1) \) is the Euler-Mascheroni constant and \( \psi(z) = \Gamma'(z)/\Gamma(z) \) is the digamma function, and the Theil index [55]

\[ T = \lim_{\theta \to 1} GE(\theta) = \frac{1}{\alpha} \left[ \psi\left(1 + \frac{1}{\alpha}\right) - \frac{1}{2} \psi\left(\frac{1}{2\kappa} - \frac{1}{2\alpha}\right) - \frac{1}{2} \psi\left(\frac{1}{2\kappa} + \frac{1}{2\alpha}\right) \right. \]

\[ - \ln{(2\kappa)} + \alpha \ln\left(\frac{\beta}{\alpha + \kappa}\right). \quad (54) \]

Expression for each index other than for the cases (53) and (54) can be derived by straightforward substitution.\(^4\)

3.1.5 Estimation and inference

Parameter estimation for the \( \kappa \)-generalized distribution can be performed using the maximum likelihood method, which yields estimators with good statistical properties [57,58]. Assuming that all sample observations \( x = \{x_1, \ldots, x_n\} \) are independent, the likelihood function is

\[ L(x; \theta) = \prod_{i=1}^{n} f(x_i; \theta)^{w_i} = \prod_{i=1}^{n} \left\{ \frac{\alpha}{\beta} \left(\frac{x_i}{\beta}\right)^{\alpha-1} \exp\left[-(x_i/\beta)^{\alpha}\right] \right\}^{w_i}, \quad (55) \]

where \( f(x_i; \theta) \) is the probability distribution function, \( \theta = \{\alpha, \beta, \kappa\} \) is the unknown parameters vector, \( w_i \) is the weight of \( i^{th} \) observation and \( n \) is the sample size. This leads to the problem of solving the partial derivatives of the log-likelihood function

\[ l(x; \theta) = \ln[L(x; \theta)] = \sum_{i=1}^{n} w_i \ln[f(x_i; \theta)] \quad (56) \]

with respect to \( \alpha, \beta \) and \( \kappa \), which corresponds to finding the solution of the following non-linear system of equations

\[ \sum_{i=1}^{n} w_i \frac{\partial}{\partial \alpha} \ln[f(x_i; \theta)] = 0, \quad (57) \]

\[ \sum_{i=1}^{n} w_i \frac{\partial}{\partial \beta} \ln[f(x_i; \theta)] = 0, \quad (58) \]

\[ \sum_{i=1}^{n} w_i \frac{\partial}{\partial \kappa} \ln[f(x_i; \theta)] = 0. \quad (59) \]

However, obtaining explicit expressions for the maximum likelihood estimators of the three parameters by resolution of the above equations is difficult, making direct analytical solutions intractable. The derivation of any of these estimators, therefore, generally involves the use of numerical optimization algorithms.\(^5\)

\(^4\) Further income inequality measures are presented in [20]. Notice that all the measures considered here are functions of the distribution moments, whose existence depends on some conditions guaranteeing the convergence of relevant integrals. As a matter of example, the Gini coefficient (51) exists provided the mean of the distribution

\[ m = \int_0^\infty xf(x; \alpha, \beta, \kappa) \, dx \]

converges, which is true when \( \frac{\alpha}{\beta} > 1 \). As shown by [56], the problem of existence of popular inequality measures is common to various parametric models of income distribution.

\(^5\) As already noted in Section 3.1.1, references [14–16] use a slightly different parametrization of the \( \kappa \)-generalized PDF, namely \( f(z; \alpha, \lambda, \kappa) = \alpha \lambda^{\alpha-1} \exp\left[-(\lambda z)^{\alpha}\right] \sqrt{1+\lambda z^{2\alpha}}, \)

where \( \lambda = \beta^{-\alpha} \) and \( z \) is defined as \( z = \frac{x}{m} \), being \( x \) the absolute income and \( m \) its mean value. Taking into account the meaning of the variable \( z \), the mean value results to be
3.2 The $\kappa$-generalized mixture model for the size distribution of wealth

The $\kappa$-generalized distribution was also successfully used in a three-component mixture model for analyzing the singularities of survey data on net wealth, i.e. the value of gross wealth minus total debt, which present highly significant frequencies of households or individuals with null and/or negative wealth [19, 20]. The support of the $\kappa$-generalized mixture model for net wealth distribution is the real line $\mathbb{R} = (-\infty, \infty)$, thus allowing to fit the subset of economic units with nil and negative net worth. Furthermore, it contains as a particular case the $\kappa$-generalized model for income distributions.

More in detail, the $\kappa$-generalized model of net wealth distribution is a mixture (or a convex combination) of an atomic and two continuous distributions. The atomic distribution concentrates its unit mass of economic agents at zero, and therefore accounts for the economic units with null net wealth. The continuous distribution accounting for the negative net wealth observations is given by a Weibull function. It has a fast left tail convergence to zero, and therefore it has finite moments of all orders. The other continuous distribution, specified as the $\kappa$-generalized model (31), accounts for the positive values of net wealth and presents a heavy right tail, thus having a small number of finite moments of positive order. This different behavior at the two tails of the distribution stems from the fact that, unlike the right tail of income and (gross or net) wealth distributions—which tend slowly to zero when income and wealth tend to infinity—the distribution of the negative values (left tail) of net wealth tends very fast to zero when the variable tends to minus infinity, since economic units face a short-term challenge of either moving out of the negative range of net wealth or bankruptcy.

3.2.1 Model specification

Formally, the $\kappa$-generalized model of net wealth distribution as a mixture of an atomic and two continuous distributions is specified as

$$f(w) = \sum_{i=1}^{3} \theta_i f_i(w), \quad -\infty < w < \infty, \quad \theta_i \geq 0, \quad \sum_{i} \theta_i = 1,$$  \hspace{1cm} (60)

where $w$ denotes the wealth variable and $\theta_i$, $i = 1, \ldots, 3$, are the mixture proportions. The two-parameter Weibull density

$$f_1(w) = \frac{s}{\lambda} \left(\frac{|w|}{\lambda}\right)^{s-1} \exp \left[ -\left(\frac{|w|}{\lambda}\right)^s \right], \quad w < 0, \quad s, \lambda > 0$$  \hspace{1cm} (61)

describes the distribution of economic units with negative net wealth, while the null net wealth observations are accounted for by a distribution that concentrates its unit mass at $w = 0$, i.e.

$$f_2(0) = 1.$$  \hspace{1cm} (62)

The other continuous distribution, $f_3(w)$, accounts for the positive values of net wealth and is specified by the three-parameter $\kappa$-generalized density

$$f_3(w) = \frac{\alpha}{\beta} \left(\frac{w}{\beta}\right)^{\alpha-1} \exp \left[ -\left(\frac{w}{\beta}\right)^\alpha \right] \sqrt{1 + \kappa^2 \left(\frac{w}{\beta}\right)^2}, \quad w > 0, \quad \alpha, \beta > 0, \quad \kappa \in [0, 1).$$  \hspace{1cm} (63)

The corresponding cumulative distribution function reads

$$F(w) = \theta_1 F_1(w) + \theta_2 F_2(w) + \theta_3 F_3(w),$$  \hspace{1cm} (64)

equal to unity, i.e. $z = \int_0^\infty zf(z; \alpha, \lambda, \kappa) \, dz = 1$. The latter relationship permits to express the scale parameter $\lambda$ as a function of the shape parameters $\alpha$ and $\kappa$, obtaining $\lambda = \frac{1}{\sqrt{\alpha \kappa}} \left(\frac{\Gamma(\frac{1}{\alpha}) \Gamma(\frac{1}{\alpha} + \frac{1}{\kappa})}{\Gamma(\frac{1}{\alpha} + \frac{1}{\kappa} + 1)}\right)^{\alpha}$. In this way, the problem to determine the values of the free parameters $\{\alpha, \lambda, \kappa\}$ from the empirical data reduces to a two-parameter $\{\alpha, \kappa\}$ fitting problem. To find the parameter values such that the negative of $l(\alpha, \theta)$ is minimized, one can use the constrained maximum likelihood estimation method [59], which solves the general maximum log-likelihood problem (56) subject to the non-linear equality constraint given by $\lambda$ and bounds $\alpha, \lambda > 0$ and $\kappa \in [0, 1)$. 

11
with

\[ F_1(w) = \begin{cases} 
\exp \left[ - \left( \frac{|w|}{\lambda} \right)^s \right] & \text{if } w < 0, \\
1 & \text{if } w \geq 0;
\end{cases} \]  

(65a)

\[ F_2(w) = \begin{cases} 
0 & \text{if } w < 0, \\
1 & \text{if } w \geq 0;
\end{cases} \]  

(65b)

\[ F_3(w) = \begin{cases} 
0 & \text{if } w \leq 0, \\
1 - \exp \left[ - \left( \frac{w}{\beta} \right)^\alpha \right] & \text{if } w > 0.
\end{cases} \]  

(65c)

It follows easily that

\[ F(w) = \begin{cases} 
\theta_1 \exp \left[ - \left( \frac{|w|}{\lambda} \right)^s \right] & \text{if } w < 0, \\
\rho & \text{if } w = 0, \\
\rho + (1 - \rho) \left( 1 - \exp \left[ - \left( \frac{w}{\beta} \right)^\alpha \right] \right) & \text{if } w > 0,
\end{cases} \]  

(66)

where \( \rho = \theta_1 + \theta_2 \) and \( 1 - \rho = \theta_3 \). It can be verified that when \( \theta_1 = \theta_2 = 0 \) (hence \( \rho = 0 \)) and \( \theta_3 = 1 - \rho = 1 \), the \( \kappa \)-generalized distribution function (30) is recovered.

3.2.2 Moments of the \( \kappa \)-generalized mixture model for net wealth distribution

From (60), the \( r \)-th order moment about the origin is

\[ E(W^r) = \int_{-\infty}^{\infty} w^r f(w) \, dw = \theta_1 E_1(W^r) + \theta_2 E_2(W^r) + \theta_3 E_3(W^r), \]  

(67)

with

\[ E_1(W^r) = (-1)^r \lambda^r \Gamma \left( 1 + \frac{r}{s} \right), \]  

(68a)

\[ E_2(W^r) = 0, \]  

(68b)

\[ E_3(W^r) = \beta^r (2\kappa)^{-\frac{r}{2}} \frac{\Gamma \left( 1 + \frac{r}{s} \right) \Gamma \left( \frac{1}{2\kappa} - \frac{r}{2\alpha} \right)}{\Gamma \left( \frac{1}{2\kappa} + \frac{r}{2\alpha} \right)}. \]  

(68c)

Specifically, the mean net wealth equals

\[ E(W) = m = -\theta_1 \lambda \Gamma \left( 1 + \frac{1}{s} \right) + \theta_3 E_3(W), \]  

(69)

where \( E_3(W) \) is given by Equation (68c) with \( r = 1 \) [compare with (45)].

3.2.3 The Lorenz curve and the Gini index of the net wealth distribution model

Given the mathematical structure of the net wealth distribution function specified by (66), we have

\[ L(u) = \begin{cases} 
-\frac{\lambda \theta_1}{m} \Gamma \left( 1 + \frac{1}{s}, \log \frac{\theta_1}{u} \right) & \text{if } 0 \leq u < \theta_1, \\
-\frac{\lambda \theta_1}{m} \Gamma \left( 1 + \frac{1}{s} \right) & \text{if } \theta_1 \leq u \leq \rho, \\
\frac{1}{m} \left[ (1 - \rho) \beta \left( \frac{1}{2\kappa} - \frac{1}{2\alpha}, 1 + \frac{1}{\alpha} \right) \right] & \text{if } u > \rho,
\end{cases} \]  

(70)

12
with \( L(1) = 1 \). Equation (70) determines the path of the Lorenz curve for the \( \kappa \)-generalized mixture model over the whole range of wealth. A graphical representation of its shape is given in Figure 5. Assuming mean wealth for the whole population is positive, as in panel (a) of the figure, the first part of \( L(u) \) drops below the horizontal line at zero and has a negative slope for negative net wealth values; then the curve is horizontal, corresponding to households that have zero net wealth, and has the conventional positive slope over the remaining units (with positive wealth values). Relative to the conventional picture of a Lorenz curve, the curve (70) takes on an even more non-standard shape when \( m < 0 \), as in this case it appears to be flipped vertically—see panel (b) of the figure. Indeed, starting from the poorest unit, the curve has a positive slope and lies above the zero line over the range of negative wealth values; then it is horizontal where wealth is zero, and has a negative slope over the remaining (positive) wealth units.

Following from the fact that the net wealth Lorenz curve presents negative values for all \( u \leq \rho \), it can be proved that the Gini index takes the general form \[ G = \frac{2}{1 + \rho |L(\theta_1)|} \int_0^1 [u - L(u)] \, du = \frac{1 - 2}{1 - \rho L(\theta_1)} \int_0^1 L(u) \, du. \] (71)

where

\[ \int_0^1 L(u) \, du = \int_0^{\theta_1} L(u) \, du + \int_{\theta_1}^{\rho} L(u) \, du + \int_{\rho}^1 L(u) \, du. \] (72)

Using (70), the Gini of the \( \kappa \)-generalized mixture model for net wealth distribution becomes

\[ G = \frac{m - 2 \left[ \frac{(1-\rho)^2 \beta}{(2\alpha)^{\frac{1}{2}}} B \left( \frac{1}{2}, \frac{1}{\alpha} \right) + 1 + \frac{1}{\alpha} \right] - \lambda \theta_1 \left( 1 - \theta_1 2^{-1-\frac{1}{\alpha}} \right) \Gamma \left( 1 + \frac{1}{\alpha} \right)}{m + \rho \lambda \theta_1 \Gamma \left( 1 + \frac{1}{\alpha} \right)}. \] (73)

Hence, while \( \beta \) and \( \lambda \) are scale parameters, all the others—i.e. \( s, \alpha, \kappa, \theta_1, \theta_2 \) and \( \theta_3 \)—are inequality parameters. Specifically, \( s, \theta_1 \) and \( \theta_2 \) account for the contribution to the Gini index of the negative and null net wealth observations, whereas \( \alpha, \kappa \) and \( \theta_3 \) account for the contribution of the positive observations.

The Gini coefficient of net wealth distribution is well-defined when households have negative or zero wealth holdings, but estimates of the coefficient might be greater than one if those households represent a relatively large fraction of the population; furthermore, whenever the mean net wealth is negative, estimates of the index are even negative [62–66].
3.3 Four-parameter extensions of the $\kappa$-generalized distribution

Quite recently, two four-parameter extensions of the $\kappa$-generalized distribution were introduced by [47] and named extended $\kappa$-generalized distributions of the first and second kind (or ExG1 and ExG2, for short).

The first kind of generalization, the ExG1, was obtained by “Weibullizing” a two-parameter deformed exponential, implicitly defined as the inverse of the following two-parameter generalization of the logarithm function [67,68]

$$\ln_{[\kappa, r]}(x) = x^r x^\kappa - x^{-\kappa},$$

which recovers the standard logarithm in the limit $(\kappa, r) \rightarrow (0, 0)$ independently of the direction.

No closed-form expression of the cumulative distribution function is available for the ExG1, while the quantile function is given by the pleasantly simple formula

$$F^{-1}(u; a, b, q, r) = b \left[ - (1 - u)^r \frac{(1 - u)^{\frac{1}{1/r}} - (1 - u)^{\frac{1}{1/q}}}{1/q} \right]^{1/r}, \quad 0 < u < 1,$$

where $a, b, q > 0$ and $r < \frac{1}{2}$. When $r = 0$ (and $a = \alpha$, $b = \beta$, $q = \frac{1}{2}$) the ExG1 quantile function is equivalent to that of the $\kappa$-generalized distribution—compare with (43).

The probability density function of the ExG1 can be expressed in terms of the cumulative probabilities $u$ as follows\(^6\)

$$f_u(u; a, b, q, r) = \frac{a^{1/r} - (1 - u)^{1/r}}{b^{1/r} - (1 - u)^{1/r}} \left[ (1 - u)^{1/q} - (1 - u)^{1/r} \right]^{-1/r}. \quad (76)$$

Since $x = F^{-1}(u; a, b, q, r)$ and $u = F(x; a, b, q, r)$ for any pair of values $(x, u)$, it follows from the definition of differentiation that

$$f_u(u; a, b, q, r) = \frac{1}{dF^{-1}(u; a, b, q, r)} = \frac{dF(x; a, b, q, r)}{dx} = f(x; a, b, q, r). \quad (77)$$

This result provides a way of plotting the density function (76). Indeed, if we let $u$ take the values, say, $u = 0.01, 0.02, \ldots, 0.99$, and plot the points $(F^{-1}(u; a, b, q, r), f_u(u; a, b, q, r))$, then we get the plot of points $(x, f(x; a, b, q, r))$, i.e. the plot of the density of $x$ given by (77). Thus plots of the ExG1 probability density function can be obtained entirely from the quantile function (75).

The ExG1 density follows a power-law behavior in both of its tails, i.e.

$$f(x; a, b, q, r) \sim c_1 x^{-a-1} \quad \text{and} \quad f(x; a, b, q, r) \sim c_2 x^{-(\frac{\kappa}{2})}, \quad (78)$$

where $c_1$ and $c_2$ are positive constants. The upper-tail power-law behavior is simply the Pareto law, whereas the power-law behavior in the lower tail of the distribution is a stylized fact that seems to be born out in actual data [70,71] and was indeed identified many years ago by [72].

Figure 6 illustrates the shape of the ExG1 density for various parameter values. In the case $a > 1$, as in panels (a) and (b), $f(x; a, b, q, r) \rightarrow 0$ when $x \rightarrow 0$ and there is always an interior mode. By contrast, if $a = 1$, $f(x; a, b, q, r) \rightarrow \frac{a}{b}$ when $x \rightarrow 0$, i.e. the ExG1 density approaches a finite positive value at the left limit; furthermore, the density is monotonically decreasing for $r \leq \frac{1}{2}$, whereas for $r > \frac{1}{2}$ it permits an interior mode—see panels (c) and (d), respectively. Finally, for $a < 1$ the ExG1 density is infinite at the left limit, i.e. $f(x; a, b, q, r) \rightarrow +\infty$ when $x \rightarrow 0$, and it can be either monotonically decreasing—as in panel (e)—or have a local maximum and minimum—as in panel (f).

The second generalization, the ExG2, was derived by extending the $\kappa$-generalized Lorenz curve

$$L(u) = I_z \left( 1 + \frac{1}{a} \right) \frac{1}{2\kappa} - \frac{1}{2\alpha}, \quad z = 1 - (1 - u)^{2\kappa} = I_u^{-1} \left( \frac{1}{2\kappa} \right) \quad (79)$$

\(^6\) A probability density function expressed in terms of $u$ was called the “density quantile function” by [69].
\( f(x; a, b, q, r) \)

(a) \( a = 3.00, b = 1.00, q = 0.60 \) and \( r = 0.30 \)

(b) \( a = 3.00, b = 1.00, q = 0.60 \) and \( r = 0.80 \)

(c) \( a = 1.00, b = 1.00, q = 0.60 \) and \( r = 0.30 \)

(d) \( a = 1.00, b = 1.00, q = 0.60 \) and \( r = 0.80 \)

(e) \( a = 0.80, b = 1.00, q = 0.60 \) and \( r = 0.30 \)

(f) \( a = 0.80, b = 1.00, q = 0.60 \) and \( r = 0.80 \)

**Figure 6.** ExG1 density for various parameter values.
in the following manner

\[ L(u) = I_z \left( p + \frac{1}{2}, q - \frac{1}{2a} \right), \quad z = I_u^{-1}(p, q), \]  

(80)

where \( a = \alpha, \ q = \frac{1}{2a} \) and \( I_u^{-1}(\cdot) \) denotes the inverse of the regularized incomplete beta function.

The cumulative distribution function of the E\( \kappa \)G2 is available in closed form. It can be expressed in terms of the incomplete beta function ratio as follows

\[ F(x; a, b, p, q) = I_z(p, q), \quad z = \left( \frac{x}{b} \right)^a \left[ \sqrt{1 + \frac{1}{4} \left( \frac{x}{b} \right)^{2a}} - \frac{1}{2} \left( \frac{x}{b} \right)^a \right], \quad x > 0, \]  

(81)

where all four parameters \( a, b, p, q \) are positive. Equation (81) implies that even the quantile function is available in closed form; it is

\[ F^{-1}(x; a, b, p, q) = b z^\frac{1}{a} (1 - z)^{-\frac{1}{2a}}, \quad z = I_u^{-1}(p, q). \]  

(82)

For \( p = 1 \) (and \( a = \alpha, \ b = \beta, \ q = \frac{1}{2} \)), it follows from (81) and (82) that the E\( \kappa \)G2 is equivalent to the \( \kappa \)-generalized distribution.

The probability density function of the E\( \kappa \)G2 has the form

\[ f(x; a, b, p, q) = \frac{a}{b B(p, q)} \left[ (1 - z)^{q + \frac{1}{2a}} - \frac{1}{2} (z^a) \right], \]  

(83)

where \( z = \left( \frac{x}{b} \right)^a \left[ \sqrt{1 + \frac{1}{4} (\frac{x}{b})^{2a}} - \frac{1}{2} (\frac{x}{b})^a \right]. \) The lower and upper tails of (83) exhibit power-law (Paretian) behavior, i.e.

\[ f(x; a, b, p, q) \sim_c c_3 x^{ap-1} \quad \text{and} \quad f(x; a, b, p, q) \sim c_4 x^{-2aq-1} \]  

(84)

for constants \( c_3 \) and \( c_4 \).

Figure 7 charts the E\( \kappa \)G2 density for various parameter values. In the case \( p > \frac{1}{a} \), as in panels (a) and (b), \( f(x; a, b, p, q) \to 0 \) when \( x \to 0 \) and the density always exhibits an interior mode. By contrast, if \( p = \frac{1}{a} \), \( f(x; a, b, p, q) \to \frac{a}{B(p, q)} \) when \( x \to 0 \), i.e. the E\( \kappa \)G2 density approaches a finite positive value at the left limit; furthermore, the density is monotonically decreasing for \( q + \frac{1}{2a} \geq 1 \), whereas for \( q + \frac{1}{2a} < \frac{1}{2} \), it permits an interior mode—see panels (c) and (d), respectively. Finally, for \( p < \frac{1}{a} \) the E\( \kappa \)G2 density is infinite at the left limit, i.e. \( f(x; a, b, p, q) \to +\infty \) when \( x \to 0 \), and it can be either monotonically decreasing—as in panel (e)—or have a local maximum and minimum—as in panel (f).

Statistical tools for describing distributions and analyzing inequality can be expressed analytically for the four-parameter variants of the \( \kappa \)-generalized distribution. References [20, 47] report formulas for the moments, the Lorenz curve, the Gini index, the coefficient of variation, the mean logarithmic deviation and the Theil index for both of the models presented. In addition, they also discuss distributional properties for the reciprocal of a random variable from the E\( \kappa \)G1 or E\( \kappa \)G2 distribution.

4 Applications of \( \kappa \)-generalized models to income and wealth data

During the last decade, there have been several applications of \( \kappa \)-generalized models to real-world data on income and wealth distribution.

The earliest investigation was completed by [14], who applied re-parametrized versions of (30) and (31) to 2001–02 household incomes for Germany, Italy, and the United Kingdom. They found an excellent agreement between the model and the whole spectrum of the empirical distributions, from the low to the high incomes, including the intermediate region for which a clear deviation was discovered when the two limiting cases of the Weibull model (40) and the pure Pareto power-law (41) were used for the same interpolation purposes.

Subsequently, the same parametrization of the \( \kappa \)-generalized distribution was fitted to Australian household incomes for 2002–03 [15] and the US family incomes for 2003 [15, 16]. Once again, the model was found to describe the entire income range extremely well; in addition, it was discovered to provide very accurate estimates of the inequality level for both the countries as evaluated by means of the Lorenz curve and the associated Gini measure of inequality.
Figure 7. ExG2 density for various parameter values.

(a) $a = 2.00$, $b = 1.00$, $p = 1.00$ and $q = 1.00$

(b) $a = 5.00$, $b = 1.00$, $p = 1.00$ and $q = 1.00$

(c) $a = 2.00$, $b = 1.00$, $p = 0.50$ and $q = 0.25$

(d) $a = 5.00$, $b = 1.00$, $p = 0.20$ and $q = 0.25$

(e) $a = 2.00$, $b = 1.00$, $p = 0.10$ and $q = 0.10$

(f) $a = 5.00$, $b = 1.00$, $p = 0.10$ and $q = 0.10$
Of special interest are papers fitting several distributions to the same data, with an eye on relative performance. From comparative studies such as [17], who considered the distribution of household income in Italy for the years 1989 to 2006, it emerges that model (31) typically outperforms its three-parameter competitors such as the Singh-Maddala [73] and Dagum type I [74] distributions, apart from the GB2 which has an extra parameter. The model was also fitted by [18] to data from other household budget surveys, namely Germany 1984–2007, Great Britain 1991–2004, and the United States 1980–2005. In a remarkable number of cases, the distribution of household income follows the $\kappa$-generalized more closely than the Singh-Maddala and Dagum type I. In particular, the fit is statistically superior in the right tail of data with respect to the other competitors in many instances. Another example of comparative study is [75], who considered US and Italian income data for the 2000s. He found the three-parameter $\kappa$-generalized model to yield better estimates of income inequality even when the goodness-of-fit is inferior to that of distributions in the GB2 family. The excellent fit of the $\kappa$-generalized distribution and its ability in providing relatively more accurate estimation of income inequality were recently confirmed in a book by [20], who utilize household income data for 45 countries selected from the most recent waves of the LIS Database (http://www.lisdatacenter.org/).

The previously mentioned works were mainly concerned with the distribution of households incomes. In an interesting contribution by [19], the $\kappa$-generalized distribution was used in a three-component mixture to model the US net wealth data for 1984–2011. Both graphical procedures and statistical methods indicate an overall good approximation of the data. The authors also highlight the relative merits of their specification with respect to finite mixture models based upon the Singh-Maddala and Dagum type I distributions for the positive values of net wealth. Similar results were recently obtained by [20] when analyzing net wealth data for 9 countries selected from the most recent waves of the LWS Database (http://www.lisdatacenter.org/).

Finally, the four-parameter models discussed in Section 3.3 were used by [47] to analyze household income/consumption data for approximately 20 countries selected from Waves IV to VI of the LIS Database. The data were grouped into 22 classes and 6 different welfare variables were considered for each country dataset. To provide a comparison with alternative four-parameter models of income distribution, the GB2 and the double Pareto-lognormal (dPlN) distribution introduced by [76] were also fitted to the same datasets and welfare variables. In all cases, parameter estimates were obtained using the maximum likelihood method for grouped data and the goodness of fit assessed using both frequency-based (FB) evaluation criteria—such as the log-likelihood value—and money-amount-based (MAB) measures, which are expected to more closely reflect the accuracy of inequality estimates—some examples of MAB measures are the square root of the sum of squared errors between the observed and estimated Lorenz curve (LRSSE) and the absolute error between the observed and estimated Gini index (AEG).

The $E\kappa G1$ is inferior to the $E\kappa G2$, $GB2$ and dPIN in terms of the log-likelihood value, but superior to the $E\kappa G2$ and $GB2$ according to the LRSSE and AEG in the overall evaluation. As for the comparison between the $E\kappa G1$ and dPIN in terms of MAB criteria, the former appears to be outperformed by the latter in terms of both the LRSSE and AEG.

No matter how the goodness-of-fit measures are combined, the pairwise comparisons between the $E\kappa G2$ and GB2 show that the former clearly outperforms the latter in the vast majority of cases. The $E\kappa G2$ is also dominant over the dPIN in terms of the log-likelihood value and the LRSSE for both the income and consumption variables, whereas the AEG tends to slightly favor the dPIN.

On the whole, the $E\kappa G2$ outperforms other four-parameter income distribution models. In particular, the $E\kappa G2$ is dominant over its counterparts in almost all cases, although slightly inferior to the dPIN when considering the accuracy of the estimated Gini index. Since the $E\kappa G2$ is close in form to the GB2, which is widely recognized as providing an excellent description of income distributions, this four-parameter variant of the $\kappa$-generalized distribution appears to drill a significant extension to the utility of parametric models for analysis of income distributions.

5 Deform to go strange: an explanation of why $\kappa$-generalized models are a good fit to income and wealth distributions

As discussed in previous section, parametric models belonging to the $\kappa$-generalized family provide a very good description over the entire income/wealth range, including the upper tail, and the inequality

---

The GB2 is a quite general family of parametric models for the size distribution of income introduced by [39] that nests most of the functional forms previously considered in the size distributions literature as special or limiting cases. In particular, both the Singh-Maddala and Dagum type I distributions are special cases of the GB2.
analysis expressed in terms of their parameters reveal very powerful. Lots of parametric models have been formalized the field of income and wealth distribution; the family of $\kappa$-generalized models, we think, is the most appropriate because it is very comfortable with the description of “strange” quantities such as income and wealth [34].

Intuitively, and generally enough, a quantity is “strange” if its probability field is not homogeneously dominated by the same statistical principle over the whole support. That is: depending on the orders of magnitude of the measurement-events, different parts of the realization field are domain of different statistical principles. Stated otherwise, a quantity is strange if below some given order of magnitude no exclusion principle is operating, while it is so above. Real sample data show that there is more difference between the richest and the second richest receivers than the poorest and the second poorest, thus suggesting that income and wealth are strange quantities: this is due to the fact that earning and saving capabilities are relative to the income and wealth order of magnitude. Quantities which are dominated by the same statistical principle independently of the part of the income/wealth support the analysis is focused on are said to be regular, whereas those dominated by different statistical principles depending on the orders of magnitude are instead said to be strange.

By considering the underlying random variable as regular, one usually comes to infer a model belonging to the family of exponential distributions. This approach works fairly well to describe the central part of the income/wealth distribution, but does a poor job in describing the lower and upper ends. To explain the tails, principally the upper tail, the empirical distribution is often broken into two pieces by estimating the leftmost point (threshold) beyond which the distribution is better described by a parametric model belonging to the family of power-law distributions. Therefore, if one privileges the wholeness of the distribution he should accept a poor fit of the upper tail, whereas if he is interested in a better interpolation of such a tail he should accept a model composed of two distinct distributions. This is fully consistent with what observed above: income and wealth are not regular quantities; rather, they behave strangely and, as it will be discussed, $\kappa$-generalized models are the most comfortable with the description and interpretation of strange quantities because they unify two distinct statistical principles into a single probability distribution by activating one of the two depending on the orders of magnitude of the realizations.

The fundamental characteristic of strange quantities is $a$-symmetry, suggesting their intrinsic relativism according to a relativistic interpretation of statistical mechanics [22–25, 29]. Moreover, the determinants of such a strange characteristic can be understood in terms of the fundamental categories of socioeconomic complex systems [34].

Three main aspects can be isolated for $a$-symmetry: structural, behavioral and formal. Structural $a$-symmetry is due to the heterogeneity of observation units populating systems at different levels of granularity. Behavioral $a$-symmetry is due to direct interaction among elementary units—i.e. in terms of transfers of portions of the quantity—and indirect interaction among sub-systems—i.e. due to the transitions of observation units from a sub-system to another while bringing with themselves their own endowments. Formal $a$-symmetry, in turn, is related to the asymmetric shape of the distribution as a representation of the system as a whole, an emergent phenomenon due to the continuous interaction of heterogeneous constituents causing inequality of the distribution. All such traits can be detected when looking at income and wealth as strange quantities.

While making inference on strange quantities, standard numeric transformations do not allow for the formalization of a unified model when two statistical principles (e.g. the exponential and the power-law families) dominate different parts of the support of a random variable. Therefore, as in relativity theory one can explain within a unified framework both the Galilean and the Einstein relativity using the Lorentz transform on the non-Euclidean Minkowski space-time, in much the same way the mathematics of deformations developed by [22–25, 29] allows for a unified explanation of both the regular and the strange quantities over their whole support.

Usually, if not aware that $X$ is a strange quantity one treats it regularly, for instance by using (alone or in combination) standard numeric transformations like scaling on the average, max-min normalizations, logarithmic transformations or root filtering. The effect of such transformations is much less effective the more the quantity is strange. This is because they operate in the same way over the whole support of $X$

---

8 The notion of “$a$-symmetry” was developed by [33] and it should not be misunderstood with that of “asymmetry”, which is included in the former. Indeed, $a$-symmetry means a lack of symmetry of the same statistical principle due to the coexistence of two distinct statistical principles over the whole support of the quantity: up to given orders of magnitude there is no exclusion principle, while beyond such a limit an exclusion principle exists. That is, according to a weak analogy with physics, it is like units following the Bose–Einstein statistics at low orders of magnitude (bosons) whereas at higher orders of magnitude they follow the Fermi–Dirac statistics (fermions).
even though they should differently react to data with respect to their orders of magnitude, which hide the absence/presence of an exclusion principle. Therefore, even after such transformations, it often happens that the rarefied tails of the inferred distribution are not well interpreted as the more dense central part. The reason resides in the fact a regular transformation—i.e. based on the regular mathematics—cannot grasp the strangeness of realizations obeying or not an exclusion principle that depends on their order of magnitude; to do this one needs a deformed mathematics.

Models belonging to the $\kappa$-generalized family are the best performing ones because, by involving a deformed mathematics, they succeed where the others fail. Model (31), for instance, behaves classically ($\kappa = 0$) when the quantity is regular, hence reconciling with the family of exponential distributions; in contrast, when the quantity is strange, it behaves relativistically ($0 < \kappa < 1$): as seen in Section 3.1.3, below a given order of magnitude the $\kappa$-generalized reconciles with a model from the exponential family of distributions, while above that it follows a Pareto power-law distribution. The interesting aspect is that (31) is a single model, activating the exponential or power-law behavior depending on the order of magnitude of the underlying quantity realizations. This, we believe, makes $\kappa$-generalized models the most appropriate ones for inference on distributions of income and wealth as strange quantities.

6 Concluding remarks

This paper has provided a brief exposition of all that appeared in the recent literature about the $\kappa$-generalized distribution and its extensions, a new and fruitful set of statistical models for the size distribution of income and wealth developed by some of us over several years of collaborative and multidisciplinary research. Given that the distribution only began to appear in the econophysics literature in 2007, it is safe to predict that there will be many further applications. On the theoretical side, especially as regards economics, there are still some unresolved issues, including the definition of a theoretical model able to demonstrate the emergence of $\kappa$-generalized income/wealth distributions as the result of decentralized interactions of a large number of heterogeneous agents. Hence, the end of the story is yet to come.

References

1. A. Chatterjee, S. Yarlagadda, B.K. Chakrabarti, Econophysics of Wealth Distributions (Springer-Verlag Italia, Milan, 2005)
2. V.M. Yakovenko, in Encyclopedia of Complexity and System Science, edited by R.A. Meyers (Springer-Verlag, Berlin, 2009), pp. 2800–2826
3. V.M. Yakovenko, J. Barkley Rosser, Jr., Rev. Mod. Phys. 81, 1703 (2009)
4. B.K. Chakrabarti, A. Chakraborti, S.R. Chakravarty, A. Chatterjee, Econophysics of Income and Wealth Distributions (Cambridge University Press, New York, 2013)
5. J.E. Stiglitz, The Price of Inequality: How Today’s Divided Society Endangers Our Future (W. W. Norton & Company, New York, 2012)
6. T. Piketty, Capital in the Twenty-First Century (The Belknap Press of Harvard University Press, Cambridge MA, 2014)
7. A.B. Atkinson, Inequality: What Can Be Done? (Harvard University Press, Cambridge MA, 2015)
8. J.E. Stiglitz, The Great Divide: Unequal Societies and What We Can Do About Them (W. W. Norton & Company, New York, 2015)
9. V. Pareto, Giorn. Econ. 10, 59 (1895)
10. V. Pareto, 1896, reprinted in Œeuvres complètes de Vilfredo Pareto, Tome 3: Écrits sur la courbe de la répartition de la richesse, edited by G. Busino (Librairie Droz, Geneva, 1965), pp. 1–15
11. V. Pareto, Cours d’économie politique (Macmillan, London, 1897a)
12. V. Pareto, Giorn. Econ. 14, 15 (1897b)
13. C. Kleiber, S. Kotz, Statistical Size Distributions in Economics and Actuarial Sciences (John Wiley & Sons, New York, 2003)
14. F. Clementi, M. Gallegati, G. Kaniadakis, Eur. Phys. J. B 57, 187 (2007)
15. F. Clementi, T. Di Matteo, M. Gallegati, G. Kaniadakis, Physica A 387, 3201 (2008)
16. F. Clementi, M. Gallegati, G. Kaniadakis, J. Stat. Mech. 2009, P02037 (2009)

To be true, the logarithmic transform usually performs well because it operates more strongly on very small/high values, but it is not enough to grasp the “strangeness” of strange quantities. On the other hand, it is not a coincidence that the $\kappa$-generalized model relies on the $\kappa$-logarithm and $\kappa$-exponential operators.
17. F. Clementi, M. Gallegati, G. Kaniadakis, Empir. Econ. 39, 559 (2010)
18. F. Clementi, M. Gallegati, G. Kaniadakis, J. Econ. 105, 63 (2012a)
19. F. Clementi, M. Gallegati, G. Kaniadakis, J. Stat. Mech. 2012, P12006 (2012b)
20. F. Clementi, M. Gallegati, *The Distribution of Income and Wealth: Parametric Modeling with the \( \kappa \)-Generalized Family* (Springer International Publishing, Switzerland, 2016)
21. G. Kaniadakis, Physica A 296, 405 (2001)
22. G. Kaniadakis, Phys. Rev. E 66, 056125 (2002)
23. G. Kaniadakis, Phys. Rev. E 72, 036108 (2005)
24. G. Kaniadakis, Eur. Phys. J. B 70, 3 (2009a)
25. G. Kaniadakis, Eur. Phys. J. A 40, 275 (2009b)
26. B. Trivellato, Math. Method. Oper. Res. 69, 1 (2009)
27. D. Imparato, B. Trivellato, in *Algebraic and Geometric Methods in Statistics*, edited by P. Gibilisco, E. Riccomagno, M.P. Rogantin, H.P. Wynn (Cambridge University Press, New York, 2010), pp. 307–26
28. B. Trivellato, Int. J. Theor. App. Finance 15, 1250038 (2012)
29. G. Kaniadakis, Entropy 15, 3983 (2013)
30. B. Trivellato, Entropy 15, 3471 (2013)
31. E. Moretto, S. Pasquali, B. Trivellato, Physica A 446, 246 (2016)
32. D. Rajoarison, D. Bolduc, H. Jayet, Econ. Lett. 86, 13 (2006)
33. D. Rajoarison, Econ. Lett. 100, 396 (2008)
34. S. Landini, in *The Distribution of Income and Wealth: Parametric Modeling with the \( \kappa \)-Generalized Family*, by F. Clementi, M. Gallegati (Springer International Publishing, Switzerland, 2016), pp. 93–102
35. R. D’Addario, Giorn. Econ. Ann. Econ. 33, 205 (1974)
36. C.P.A. Bartels, H. van Metelen, *Alternative probability density functions of income: A comparison of the lognormal-, Gamma- and Weibull-distribution with Dutch data*. Research Memorandum No. 29 (1975)
37. C.P.A. Bartels, *Economic aspects of regional welfare, income distribution and unemployment* (Martinus Nijhoff, Leiden, 1977)
38. P. Espinguet, M. Terraza, Econ. Appl. 36, 535 (1983)
39. J.B. McDonald, Econometrica 52, 647 (1984).
40. N. Atoda, T. Suruga, T. Tachibanaki, Econ. Stud. Quart. 39, 14 (1988)
41. R.F. Bordley, J.B. McDonald, A. Mantrala, J. Income Distrib. 6, 91 (1996)
42. K. Brachmann, A. Stich, Trede, All. Stat. Arch. 80, 285 (1996)
43. T. Tachibanaki, T. Suruga, N. Atoda, J. Japan. Statist. Soc. 27, 191 (1997)
44. B. Mandelbrot, Int. Econ. Rev. 1, 79 (1960)
45. M.O. Lorenz, Publ. Am. Stat. Assoc. 9, 209 (1905)
46. C. Gini, Atti Reale Istit. Veneto Sci. Lett. Art. 73, 1203 (1914)
47. M. Okamoto, *Extension of the \( \kappa \)-generalized distribution: New four-parameter models for the size distribution of income and consumption*. LIS Working Paper No. 600 (2013), available at: http://www.lisdatacenter.org/wps/liswps/600.pdf
48. A.A. Drăgulescu, V.M. Yakovenko, Eur. Phys. J. B 20, 558 (2001)
49. P.D. Allison, Am. Sociol. Rev. 43, 865 (1978)
50. F.A. Cowell, Eur. Econ. Rev. 13, 147 (1980a)
51. F.A. Cowell, Rev. Econ. Stud. 47, 521 (1980b)
52. A.F. Shorrocks, Econometrica 48, 613 (1980)
53. F.A. Cowell, K. Kuga, J. Econ. Theory 25, 131 (1981a)
54. F.A. Cowell, K. Kuga, Eur. Econ. Rev. 15, 287 (1981b)
55. H. Theil, *Economics and Information Theory* (North-Holland, Amsterdam, 1967)
56. C. Kleiber, Econ. Lett. 57, 39 (1997)
57. C. Rao, *Linear Statistical Inference and its Applications* (John Wiley & Sons, New York, 1973)
58. J.K. Ghosh, *Higher Order Asymptotics* (Institute of Mathematical Statistics and American Statistical Association, Hayward CA, 1994)
59. R.J. Schoenberg, Comput. Econ. 10, 251 (1997)
60. C. Dagum, in *Encyclopedia of Statistical Sciences, Second Edition, Volume 5*, edited by N. Balakrishnan, C.B. Read, B. Vidakovic (John Wiley & Sons, New York, 2006a), pp. 3363–78
61. C. Dagum, Statistica 66, 235 (2006b)
62. J.B. Hagerbaumer, Rev. Econ. Stat. 59, 377 (1977)
63. G. Pyatt, C.-N. Chen, J. Fei, Q. J. Econ. 95, 451 (1980)
64. Y. Amiel, F.A. Cowell, A. Polovin, Economica 63, S63 (1996)
65. S.P. Jenkins, M. Jäntti, *Methods for summarizing and comparing wealth distributions*. ISER
66. F.A. Cowell, *Measuring Inequality* (Oxford University Press, New York, 2011)
67. G. Kaniadakis, M. Lissia, A.M. Scarfone, Physica A 340, 41 (2004)
68. G. Kaniadakis, M. Lissia, A.M. Scarfone, Phys. Rev. E 71, 046128 (2005)
69. E. Parzen, J. Am. Stat. Assoc. 74, 105 (1979)
70. W.J. Reed, Physica A 319, 469 (2003)
71. W.J. Reed, J. Income Distrib. 13, 7 (2004)
72. D.G. Champernowne, Econ. J. 63, 318 (1953)
73. S.K. Singh, G.S. Maddala, Econometrica 44, 963 (1976)
74. C. Dagum, Econ. Appl. 30, 413 (1977)
75. M. Okamoto, Econ. Bull. 32, 2969 (2012)
76. W.J. Reed, M. Jorgensen, Commun. Stat.-Theor. M. 33, 1733 (2004)