Empir Econ (2010) 39:559–591
DOI 10.1007/s00181-009-0318-2

ORIGINAL PAPER

A model of personal income distribution with application to Italian data

Fabio Clementi · Mauro Gallegati · Giorgio Kaniadakis

Received: 23 September 2008 / Accepted: 19 August 2009 / Published online: 12 November 2009

Abstract This article proposes the κ-generalized distribution as a descriptive model for the distribution and dispersion of income within a population based on the deformed exponential and logarithm functions recently introduced by Kaniadakis (Phys A 296:405–425, 2001; Phys Rev E 66:056125, 2002; Phys Rev E 72:036108, 2005). Expressions are reported which facilitate the analysis of the associated moments and various tools for the measurement of inequality. An empirical application, including a comparison of alternative distributions, is made to household income data in Italy for the years 1989 to 2006.

Keywords Income distribution · Income inequality · κ-generalized distribution

JEL Classification C16 · D31

1 Introduction

Fitting a parametric model to income data can be a valuable and informative tool for the distributional analysis. Not only one can summarize the information contained in thousands of observations, but also useful information can be drawn directly from the estimated parameters. For example, one could be interested in measuring income inequality, comparing different distributions or elaborating income redistribution

F. Clementi (✉) · M. Gallegati
Department of Economics, Polytechnic University of Marche, Piazzale R. Martelli 8, 60121 Ancona, Italy
e-mail: fabio.clementi@univpm.it

G. Kaniadakis
Department of Physics, Polytechnic University of Turin, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

policies: these concepts may sometimes be directly derived from the parameters of a fitted distribution.

The symmetric bell-shaped curve that describes the Gaussian distribution is inappropriate for describing most income distributions because they tend to be skewed with a peak in the lower-middle income range and to have a long right-hand tail. To capture these features, a large number of functional forms other than the Gaussian one have been suggested.[1] According to Dagum (1977), the different approaches can be grouped into three categories. One approach consists in viewing the functional form describing an income distribution as the outcome of a stochastic process [e.g. the lognormal model in Gibrat (1931) and the Pareto distribution in Champernowne, (1953)]. Another approach derives flexible analytical forms by considering solely their ability to ensure a satisfactory fit to empirical data [e.g. the gamma density of Salem and Mount (1974)]. Finally, models are also derived from differential equations specified to capture regularity features of observed income distributions [e.g. the models proposed by Singh and Maddala (1976) and Dagum (1977)].

The functional forms most frequently used in applied work are the two-parameter Pareto, lognormal and gamma distributions. The Pareto distribution (e.g. Arnold 1983) accurately models high levels of income, but does a poor job in describing the lower end of the distribution. If one considers the entire range of income, the fit is better for the lognormal, but the performance of the model in the upper end is far from being satisfactory (Aitchison and Brown 1954, 1957). In terms of goodness-of-fit, the gamma distribution outperforms the lognormal at the two tails of the distribution (McDonald and Ransom 1979), even though in the middle income range it overcorrects for the positive skewness of the data (Majumder and Chakravarty 1990). Better fits are obtained using three- or four-parameter models [e.g. the Singh-Maddala, Dagum and the generalized beta II distribution of McDonald (1984) and its associates] that deserve further attention for their ability to characterize cyclical movements in the observed income distribution (Metcalf 1972).

In this article, we introduce an additional three-parameter distribution that is a generalization of the Weibull distribution by exploiting recent developments on the use of deformed exponential and logarithm functions as relationships that are more flexible than the standard ones to build statistical models [see e.g. Rajaonarison et al. (2005) and Rajaonarison (2008) for an application to the context of choice modeling]. The focus is on the κ-deformed exponential and logarithm functions proposed by Kaniadakis (2001, 2002, 2005) and defined as

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

$$\ln_{\kappa}(x) = \frac{x^{\kappa} - x^{-\kappa}}{2\kappa}, \quad x \in \mathbb{R}_{+}, \tag{1b}$$

which reduce to the standard exponential and logarithm, respectively, as the deformation parameter κ approaches zero. The above functions have many very interesting properties, some being identical to the ones of the undeformed functions; for

[1] See e.g. the comprehensive monograph by Kleiber and Kotz (2003).

applications to statistical analysis of income distribution, the most interesting property is their power-law asymptotic behavior

$$\exp_\kappa(x) \underset{x\to\pm\infty}{\sim} |2\kappa x|^{\pm\frac{1}{|\kappa|}},$$
$$\ln_\kappa(x) \underset{x\to 0^+}{\sim} -\frac{1}{2|\kappa|}x^{-|\kappa|},$$
$$\ln_\kappa(x) \underset{x\to +\infty}{\sim} \frac{1}{2|\kappa|}x^{|\kappa|}.$$

Formally, the distribution presented here can be obtained by maximizing according to Jaynes (1957a,b) maximum entropy principle the Shannon (1948) information measure

$$S \equiv -\int_0^\infty f(x)\ln f(x)\,\mathrm{d}x \tag{2}$$

under the natural constraint that normalizes the density,

$$\int_0^\infty f(x)\,\mathrm{d}x = 1, \tag{3}$$

and the three characterizing moments

$$\int_0^\infty \ln x f(x)\,\mathrm{d}x = \ln\beta - \frac{1}{\alpha}\left[\gamma + \psi\left(\frac{1}{2\kappa}\right) + \ln(2\kappa) + \kappa\right], \tag{4a}$$

$$\int_0^\infty \ln\left[1+\kappa^2\left(\frac{x}{\beta}\right)^{2\alpha}\right] f(x)\,\mathrm{d}x = 2\kappa - \psi\left(1+\frac{1}{4\kappa}\right) + \psi\left(\frac{1}{2}+\frac{1}{4\kappa}\right), \tag{4b}$$

$$\int_0^\infty \ln\left[\sqrt{1+\kappa^2\left(\frac{x}{\beta}\right)^{2\alpha}} - \kappa\left(\frac{x}{\beta}\right)^{\alpha}\right] f(x)\,\mathrm{d}x$$
$$= \int_0^\infty \sinh^{-1}\left[-\kappa\left(\frac{x}{\beta}\right)^{\alpha}\right] f(x)\,\mathrm{d}x = \kappa, \tag{4c}$$

where the latter reproduces the deformation parameter κ. We show in Appendix A that the solution to the variational problem (2–4) is given by what we call the κ-generalized distribution

$$f(x;\alpha,\beta,\kappa)=\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}\frac{\exp_{\kappa}\left[-(x/\beta)^{\alpha}\right]}{\sqrt{1+\kappa^{2}(x/\beta)^{2\alpha}}},\quad x\geq 0,\quad \alpha,\beta>0,\quad \kappa\in[0,1)\,. \tag{5}$$

Virtually, all income distributions can be derived from the maximum entropy principle.[2] For example, Ord et al. (1981) and Leipnik (1990) pointed out that several well-known distributions, such as Pareto, lognormal, gamma, Singh-Maddala and generalized beta II, might be selected if one uses a criterion of maximum entropy. Nonetheless, most of the densities used in the literature are characterized by just two moment functions that take care of lower, middle and upper income levels at the same time. As it will become clear later in this article, these two moment functions are not enough to extract all the relevant information from the data, whereas adding one more moment function to the maximum entropy problem makes the specification under consideration more consistent with the amount of information (or uncertainty) present in the data, while achieving almost identical goodness-of-fit.

The rest of this article is organized as follows: in Sect. 2, we derive the basic statistical properties of the new distribution and give formulas for the shape, moments and standard tools for inequality measurement. In Sect. 3, an empirical application is made to personal income data for Italian households from 1989 to 2006: first, an attempt is made to determine how well the proposed statistical distribution fits the income data; second, the fit is checked against the performance of other existing distributions. A summary of the article is given in Sect. 4.

2 The κ-generalized distribution

2.1 Definitions and basic properties

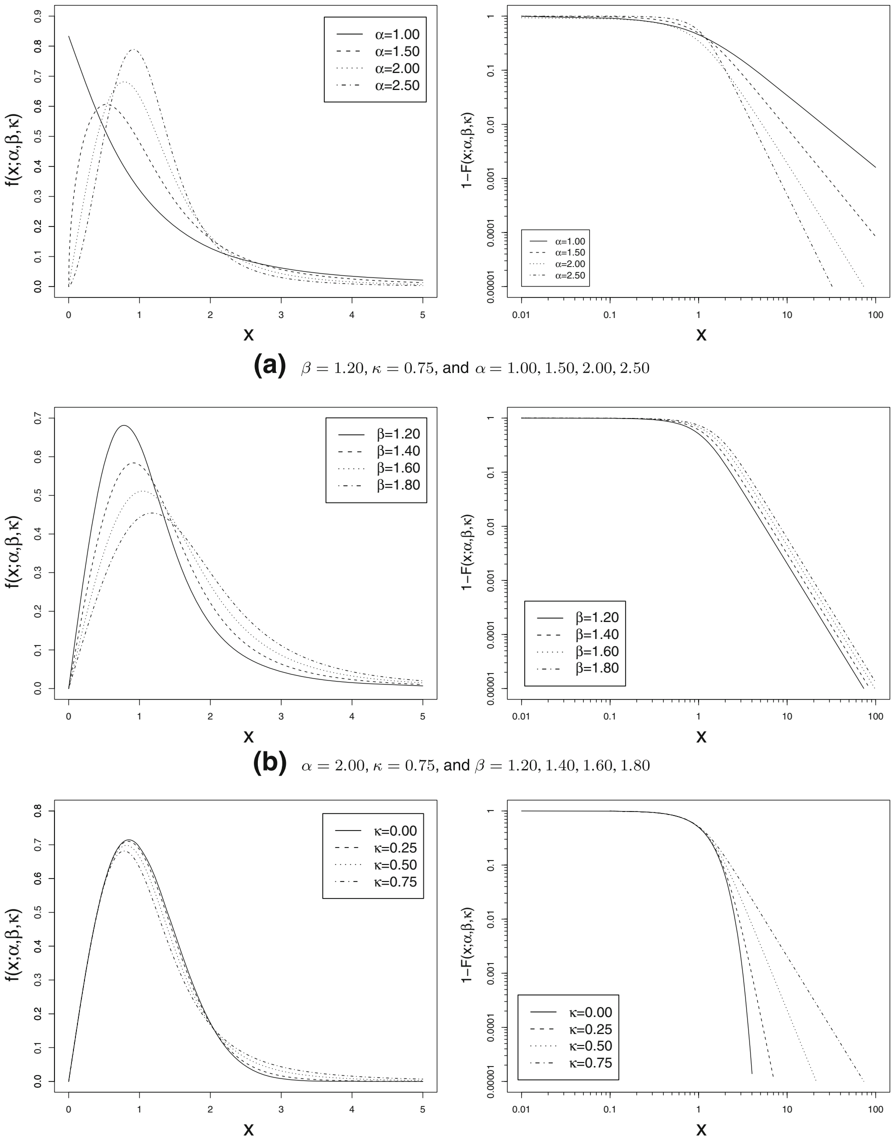
A random variable X is said to have a κ-generalized distribution, and we write $X\sim$ κ-gen (α,β,κ), if it has a probability density function given by (5). Its cumulative distribution function can be expressed as

$$F(x;\alpha,\beta,\kappa)=1-\exp_{\kappa}\left[-(x/\beta)^{\alpha}\right]. \tag{6}$$

Figure 1 illustrates the effects on the shape of the distribution of different values of the parameters.

The exponent α quantifies the curvature (shape) of the distribution, which is less (more) pronounced for lower (higher) values of the parameter. The constant β is a characteristic scale, since its value determines the scale of the probability distribution: if β is small, then the distribution will be more concentrated around the mode; if β is large, then it will be more spread out. Finally, the parameter κ measures the heaviness of the right tail: the larger (smaller) its magnitude, the fatter (thinner) the tail.

[2] See e.g. Kapur (1989) for a comprehensive account.



(a) $\beta = 1.20, \kappa = 0.75$, and $\alpha = 1.00, 1.50, 2.00, 2.50$

(b) $\alpha = 2.00, \kappa = 0.75$, and $\beta = 1.20, 1.40, 1.60, 1.80$

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

Fig. 1 κ-generalized densities (*left*) and log–log complementary distributions (*right*) for some different values of the parameters

As $\kappa \to 0$, the distribution tends to the Weibull distribution; it can be easily verified that

$$\lim_{\kappa \to 0} f(x; \alpha, \beta, \kappa) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left[-(x/\beta)^{\alpha}\right] \tag{7a}$$

and

$$\lim_{\kappa \to 0} F(x; \alpha, \beta, \kappa) = 1 - \exp\left[-(x/\beta)^{\alpha}\right]. \tag{7b}$$

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 exponential law is also a special limiting case.[3]

For $x \to 0^+$, the distribution behaves similarly to the Weibull model (7), whereas for large x it approaches a Pareto distribution with scale $\beta (2\kappa)^{-\frac{1}{\alpha}}$ and shape $\frac{\alpha}{\kappa}$, i.e.

$$f(x; \alpha, \beta, \kappa) \underset{x \to +\infty}{\sim} \frac{\frac{\alpha}{\kappa}\left[\beta (2\kappa)^{-\frac{1}{\alpha}}\right]^{\frac{\alpha}{\kappa}}}{x^{\frac{\alpha}{\kappa}+1}}$$

and

$$F(x; \alpha, \beta, \kappa) \underset{x \to +\infty}{\sim} 1 - \left[\frac{\beta (2\kappa)^{-\frac{1}{\alpha}}}{x}\right]^{\frac{\alpha}{\kappa}},$$

thus satisfying the weak Pareto law (Mandelbrot 1960).[4]

Equation 6 implies that the quantile function is available in closed form

$$F^{-1}(u; \alpha, \beta, \kappa) = \beta \left[\ln_{\kappa}\left(\frac{1}{1-u}\right)\right]^{\frac{1}{\alpha}}, \quad 0 < u < 1, \tag{8}$$

an attractive feature for the derivation of Lorenz-ordering results and simulation purposes (see Sect. 2.3). From (8), we easily determine that the median of the distribution is

$$x_{\text{med}} = \beta \left[\ln_{\kappa}(2)\right]^{\frac{1}{\alpha}}.$$

[3] The Weibull distribution was used only sporadically as an income distribution. Some quite recent applications can be found in Bartels (1977), Espinguet and Terraza (1983), McDonald (1984), Atoda et al. (1988), Bordley et al. (1996), Brachmann et al. (1996) and Tachibanaki et al. (1997).

[4] Further generalizations of the Pareto law were introduced by Kakwani (1980), $\lim_{x \to +\infty} \frac{xf(x)}{1-F(x)} = \alpha$, and Esteban (1986), $\lim_{x \to +\infty} \left[1 + \frac{xf'(x)}{f(x)}\right] = -\alpha$. Since we have

$$\lim_{x \to +\infty} \frac{xf(x; \alpha, \beta, \kappa)}{1 - F(x; \alpha, \beta, \kappa)} = \frac{\alpha}{\kappa} \quad \text{and} \quad \lim_{x \to +\infty} \left[1 + \frac{xf'(x; \alpha, \beta, \kappa)}{f(x; \alpha, \beta, \kappa)}\right] = -\frac{\alpha}{\kappa},$$

the distribution also obeys these alternative versions of the weak Pareto law.

The mode occurs at

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

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

2.2 Moments and related statistics

The rth-order moment about the origin of the κ-generalized distribution equals

$$\mu'_r = \int_0^\infty x^r f(x; \alpha, \beta, \kappa)\, \mathrm{d}x = \beta^r (2\kappa)^{-\frac{r}{\alpha}} \frac{\Gamma\left(1 + \frac{r}{\alpha}\right)}{1 + \frac{r}{\alpha}\kappa} \frac{\Gamma\left(\frac{1}{2\kappa} - \frac{r}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa} + \frac{r}{2\alpha}\right)}, \tag{9}$$

where $\Gamma(\cdot)$ denotes the gamma function, and exists for $-\alpha < r < \frac{\alpha}{\kappa}$. Specifically, $\mu'_1 = m$ is the mean of the distribution and

$$\begin{aligned}\sigma^2 &= \mu'_2 - m^2 \\ &= \beta^2 (2\kappa)^{-\frac{2}{\alpha}} \left\{ \frac{\Gamma\left(1 + \frac{2}{\alpha}\right)}{1 + 2\frac{\kappa}{\alpha}} \frac{\Gamma\left(\frac{1}{2\kappa} - \frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa} + \frac{1}{\alpha}\right)} - \left[\frac{\Gamma\left(1 + \frac{1}{\alpha}\right)}{1 + \frac{\kappa}{\alpha}} \frac{\Gamma\left(\frac{1}{2\kappa} - \frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa} + \frac{1}{2\alpha}\right)} \right]^2 \right\}.\end{aligned}$$

is the variance. Hence, the coefficient of variation is given by

$$\mathrm{CV} = \frac{\sigma}{m} = \sqrt{\frac{\frac{\Gamma\left(1+\frac{2}{\alpha}\right)}{1+2\frac{\kappa}{\alpha}} \frac{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{\alpha}\right)}}{\left[\frac{\Gamma\left(1+\frac{1}{\alpha}\right)}{1+\frac{\kappa}{\alpha}} \frac{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right)}\right]^2} - 1}.$$

It is also possible to define the standardized measures of skewness and kurtosis, respectively, given by

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu'_3 - 3\mu'_2 m + 2m^3}{\sigma^3}$$

and

$$\gamma_2 = \frac{\mu_4}{\sigma^4} = \frac{\mu'_4 - 4\mu'_3 m - 6\mu'_2 m^2 - 3m^4}{\sigma^4},$$

where $\mu_r = \sum_{j=0}^{r} \binom{r}{j} (-1)^{r-j} \mu'_j m^{r-j}$ is the moment about the mean.

2.3 Lorenz curve and inequality measures

Since the quantile function of the κ-generalized distribution is available in closed form, its normalized integral, the Lorenz (1905) curve

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

can be expressed analytically (Gastwirth 1971). Therefore, we have

$$L(u) = 1 - \frac{1+\frac{\kappa}{\alpha}}{2\Gamma\left(\frac{1}{\alpha}\right)} \frac{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right)} \left\{2\alpha\,(2\kappa)^{\frac{1}{\alpha}}\,(1-u)\left[\ln_\kappa\left(\frac{1}{1-u}\right)\right]^{\frac{1}{\alpha}} + B_X\left(\frac{1}{2\kappa}-\frac{1}{2\alpha},\frac{1}{\alpha}\right)\right.$$
$$\left. + B_X\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}+1,\frac{1}{\alpha}\right)\right\}, \tag{10}$$

where $B_X(\cdot,\cdot)$ is the incomplete beta function with $X = (1-u)^{2\kappa}$. Clearly, the curve exists if and only if $\frac{\alpha}{\kappa} > 1$. Given two κ-generalized distributions X_1 and X_2, the following theorem gives the parameter constellations for which their Lorenz curves do not intersect, and we have $X_1 \leq_L X_2$, i.e. when the Lorenz curve of X_1 lies nowhere below that of X_2 (in symbols: $L_{X_1}(u) \geq L_{X_2}(u)$, for all $u \in [0,1]$), and consequently X_1 exhibits less inequality than X_2 in the Lorenz sense:

Theorem 1 *Let* $X_i \sim \kappa\text{-gen}(\alpha_i,\beta_i,\kappa_i)$, $i = 1,2$, *be* κ*-generalized distributions. Then,*

$$X_1 \leq_L X_2 \Longleftarrow: \alpha_1 \geq \alpha_2 \quad \textit{and} \quad \frac{\alpha_1}{\kappa_1} \geq \frac{\alpha_2}{\kappa_2}. \tag{11}$$

Proof See Appendix B. □

Figure 2 provides an illustration of (11), showing that the less unequal distribution (in the Lorenz sense) always exhibits lighter tails.[5]

According to the results of Sect. 2.2, several measures of inequality can be considered. In particular, the Gini (1914) coefficient can be derived using the representation in terms of order statistics $G = 1 - m^{-1}\int_0^\infty [1-F(x)]^2\,\mathrm{d}x$ due to Arnold and Laguna (1977); this yields

$$G = 1 - \frac{2\alpha+2\kappa}{2\alpha+\kappa} \frac{\Gamma\left(\frac{1}{\kappa}-\frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{\kappa}+\frac{1}{2\alpha}\right)} \frac{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right)}. \tag{12}$$

[5] Within three-parameter families, the Lorenz-ordering has been settled by Taillie (1981) and Wilfling (1996b) for the generalized gamma distribution, and by Wilfling and Krämer (1993) and Kleiber (1996) for the most popular Singh-Maddala and Dagum type I distributions, respectively. Both these distributions are special cases of McDonald (1984) four-parameter generalized beta II distribution. For this family, Wilfling (1996a) was able to derive a necessary and also a sufficient condition.

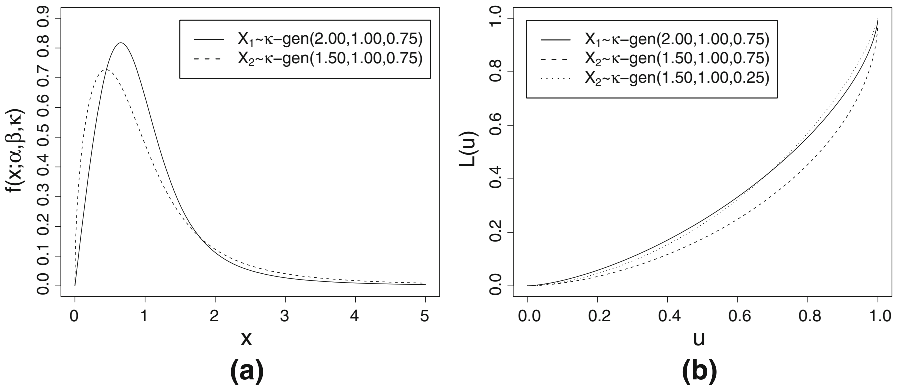


Fig. 2 Tails **a** and Lorenz curves **b** for two κ-generalized distributions. As shown by the dotted line, the Lorenz curves intersect if the parameters are not chosen according to Theorem 1

Furthermore, the generalized entropy class of inequality measures (Cowell 1980a,b; Shorrocks 1980; Cowell and Kuga 1981a,b) is given by

$$GE\,(\theta) = \frac{1}{\theta^2 - \theta}\left\{\left(\frac{\beta}{m}\right)^{\theta}\left[\frac{(2\kappa)^{-\frac{\theta}{\alpha}}}{1+\frac{\theta}{\alpha}\kappa}\frac{\Gamma\left(\frac{1}{2\kappa}-\frac{\theta}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{\theta}{2\alpha}\right)}\Gamma\left(1+\frac{\theta}{\alpha}\right)\right]-1\right\},\quad \theta \neq 0, 1. \tag{13}$$

Equation 13 defines a class because the index $GE\,(\theta)$ assumes different forms depending on the value assigned to θ. In applied work, two limiting cases of (13) are of particular interest for inequality measurement: the mean logarithmic deviation index

$$MLD = \lim_{\theta\to 0} GE_{\kappa}\,(\theta) = \frac{1}{\alpha}\left[\gamma + \psi\left(\frac{1}{2\kappa}\right) + \ln(2\kappa) - \alpha \ln\left(\frac{\beta}{m}\right) + \kappa\right], \tag{14}$$

where $\gamma = -\psi\,(1)$ is the Euler–Mascheroni constant and $\psi\,(z) = \Gamma'\,(z)\,/\,\Gamma\,(z)$ is the digamma function, and the Theil (1967) index

$$T = \lim\nolimits_{\theta\to 1} GE_{\kappa}\,(\theta) = \tfrac{1}{\alpha}\Big[\psi\left(1+\tfrac{1}{\alpha}\right) - \tfrac{1}{2}\psi\left(\tfrac{1}{2\kappa}-\tfrac{1}{2\alpha}\right) - \tfrac{1}{2}\psi\left(\tfrac{1}{2\kappa}+\tfrac{1}{2\alpha}\right) - \ln(2\kappa) + \alpha \ln\left(\tfrac{\beta}{m}\right) - \tfrac{\alpha\kappa}{\alpha+\kappa}\Big]. \tag{15}$$

Expression for each index other than for the cases (14) and (15) can be derived by straightforward substitution. In particular, the Atkinson (1970) class of inequality indices can be easily computed from (13) by exploiting the relationship [e.g. Cowell (1995)]

$$A\,(\epsilon) = 1 - \left[\epsilon\,(\epsilon-1)\,GE\,(1-\epsilon)+1\right]^{\frac{1}{1-\epsilon}}, \tag{16}$$

where $\epsilon = 1 - \theta$, $\epsilon > 0$ and $\epsilon \neq 1$, is the inequality aversion parameter. The limiting form of (16) as $\epsilon \to 1$ is

$$A(1) = 1 - \exp(-MLD).$$

3 Empirical results

The proposed statistical distribution has been fitted to Italian household incomes for the years 1989 to 2006. The source of data is the Annual database of the Bank of Italy "Survey on Household Income and Wealth" (SHIW), including nine biennial waves of data on around 8,000 households (the survey for 1997 was shifted to 1998).[6] Income in the SHIW is recorded net of payments of taxes and social security contributions. It is the sum of four main components: compensation of employees, pensions and net transfers, net income from self-employment, property income (including income from buildings and income from financial assets). We have omitted from the analysis observations with zero or negative values for income[7] and converted, when necessary, to euros (i.e. for the period 1989–2000). Furthermore, incomes have been equivalized for differences in household size[8] and weighted by using appropriate sampling weights provided by the Bank of Italy. Finally, we have deflated all figures so as to obtain distributions of "real" income. To do so, we have employed the consumer price index deflator (yearly series based on year 2000) reported by the OECD.[9]

Table 1 contains several summary statistics of the distribution of Italian household income for the period considered.

According to the SHIW evidence, the fall in household incomes caused by the recession of the early 1990s continued for some time. In 1995, real mean income was still around 3% below its 1991 level. It then rose in each of the following years, reaching about the same level of 1989 only in 2002. As displayed in Fig. 3a, the pattern of evolution of Gini and Theil coefficients over the sample period indicates that the negative business cycle led to a sharp rise in inequality between 1991 and 1993, followed by a substantial stability in the following years, save for a temporary increase in 1998 and

[6] The SHIW data are publicly available online at http://www.bancaditalia.it/statistiche/indcamp/bilfait. For the main features of the SHIW, its sample design, interviewing procedures and response rates, see Brandolini (1999).

[7] The possibility that even a few incomes may be nonpositive raises some issues of principle for inequality measurement. For example, many of the standard inequality measures are simply undefined for negative incomes and some of these measures will not work even for zero incomes [see e.g. Amiel et al. (1996)]. For these reasons, we adopt here the common practice in the formal literature on income inequality and measurement of taking only positive values for income.

[8] In this article, the "modified OECD" equivalence scale has been used. This scale allocates points to each person in a household by taking the first adult as having a weight of 1 point, whereas each additional person who is 14 years or older is allocated 0.5 points, and each child under the age of 14 is allocated 0.3 points. Equivalized household income is derived by dividing total household income by a factor equal to the sum of the equivalence points allocated to the household members. The modified OECD scale enjoys a wide degree of acceptance. Since the late 1990s, it has been adopted as a standard by Eurostat for all statistics on income and living conditions, by replacing the old OECD scale. Unlike the old scale, the modified one gives less weight to any additional household member, allowing for higher economies of scale.

[9] Available at http://www.sourceoecd.org.

Table 1 Standard summary statistics of household income in Italy during the period 1989–2006

Stats	Wave								
	1989	1991	1993	1995	1998	2000	2002	2004	2006
Obs	8,261.0000	8,185.0000	8,066.0000	8,122.0000	7,112.0000	7,958.0000	7,990.0000	8,008.0000	7,762.0000
Min	23.1266	89.7680	6.4328	16.0258	6.0508	7.4886	4.9509	118.0191	5.1127
1st quartile	8,859.6290	8,343.6910	7,390.9830	7,589.5900	8,053.6150	8,350.6780	8,477.3100	8,531.0920	9,275.4720
Median	12,715.1800	12,038.6400	11,335.4000	11,359.8400	12,165.5700	12,537.5100	12,809.8500	13,068.1900	13,741.1600
3rd quartile	17,858.6300	16,978.2500	16,530.0400	16,443.8500	17,125.0900	17,707.0900	17,968.1200	18,578.8200	19,514.5600
Max	232,808.2000	175,548.7000	186,633.8000	198,932.1000	385,231.0000	336,017.2000	245,387.8000	657,256.0000	704,859.2000
Mean	14,709.0100	13,659.8200	13,184.9400	13,260.2600	14,201.5500	14,492.7100	14,630.8500	15,196.7600	15,929.2300
Std. dev.	9,835.6910	8,288.0260	9,059.5000	9,526.0020	11,751.0700	10,704.4900	10,244.9000	12,847.0600	13,555.4000
Skewness	5.4675	3.9879	2.9714	4.3477	7.9172	6.2204	4.7936	12.1035	14.8546
Kurtosis	78.8876	49.4549	24.1691	49.3531	152.6254	108.3715	65.8369	365.6869	551.4019
Gini	0.3020	0.2913	0.3338	0.3307	0.3401	0.3271	0.3217	0.3306	0.3231
Theil	0.1617	0.1462	0.1919	0.1934	0.2183	0.1932	0.1839	0.2076	0.2004
	Quintile group share (%)								
1st	8.3200	8.3800	6.8600	7.0400	6.7200	7.1800	7.3600	7.3200	7.4400
2nd	13.0100	13.3400	12.3200	12.5200	12.5100	12.5800	12.7100	12.3900	12.7200
3rd	17.2800	17.6500	17.2500	17.2100	17.1900	17.3700	17.4400	17.1900	17.2700
4th	22.9900	23.2000	23.3900	23.1100	22.7000	22.9900	22.8600	22.7700	22.7800
5th	38.4100	37.4300	40.1800	40.1300	40.8800	39.8900	39.6400	40.3200	39.8000

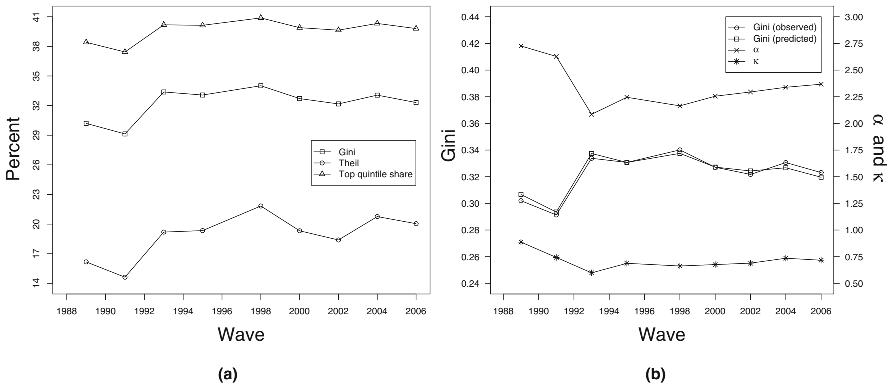


Fig. 3 Inequality **a** and the κ-generalized distribution **b** across years

2004. Looking at the quintile income shares, we observe that the households in the top quintile hold about two-fifths of the total income, while the two-fifths of households with the lowest income (i.e. the bottom 40%) receive about one-fifth of the overall income. The share of the total income received by the bottom 20% of households is just about 7% over the whole period. In particular, the evolution of the income share of top 20% seems to have driven the inequality in income distribution: after rising to a peak in 1998, the income share of top quintile went down and then it increased again in 2004.

The maximum likelihood estimates of the model for all the years, including the negative log-likelihood value ($-\ln L$) as well as the values of Akaike (1973) and Schwarz (1978) information criteria (AIC and BIC), are presented in Tables 2, 3, 4.[10] The parameters were very precisely estimated, and the fit, as judged by visual inspection of Figs. 4, 5, 6, was fairly good over the whole range of income for all sets of data. In particular, the scale parameter (β) reflects the changes in real mean income over the period. The other parameters (α and κ), characterizing distributional shape, are easiest to interpret by comparing predicted values for key distributional summary measures with their sample counterparts, as the effect of changing one of them is contingent on the value of the other parameter. For example, Fig. 3b shows that the Gini implied by the estimated functional form tracks the statistic calculated from the data closely—the correlation coefficient between the two series is 0.9821, which is highly significant (p value < 0.0001). The changes of the parameters α and κ are the mirror images of the pattern of inequality over time. Therefore, to give some idea of how inequality depends on the two parameters, and to assess overall goodness-of-fit, we plot in Fig. 7 the Gini coefficient given by Eq. 12 as a function of α and κ (the range of values for the parameters is equal to that estimated during the sample period).

As suggested by the positively sloped contours, increases in α and κ are associated with declining levels of the Gini coefficient, whereas the profile of inequality is reversed if one moves along the opposite direction of change. Furthermore, higher

[10] Likelihoods have been maximized with respect to the parameters by using a PORT routine as supplied by the R function `nlminb` (R Development Core Team 2008).

Table 2 Estimated distribution functions for the Italian household income in 1989, 1991 and 1993[a]

Wave	Distribution	Parameters[b]				$-\ln L$[c]	AIC	BIC	Gini[d]
		a ($\boldsymbol{\alpha}$)	b ($\boldsymbol{\beta}$)	p ($\boldsymbol{\kappa}$)	q				
	κ-gen	**2.7258**	**14,300.1274**	**0.8868**	–	252,026.2000	504,058.3000	504,079.4000	0.3067
		(0.0244)	**(63.4375)**	**(0.0195)**					
	SM	3.1292	13,303.3673	–	1.1152	251,890.3000	503,786.6000	503,807.7000	0.3032
1989									
		(0.0311)	(173.0173)		(0.0290)	(86.0735)**			
	D	3.2368	12,591.1727	1.0069	–	251,899.4000	503,804.7000	503,825.8000	0.3085
		(0.0333)	(156.3360)	(0.0256)		(104.1777)**			
	GB2	2.1893	13,111.5984	1.7717	1.8910	251,847.3000	503,702.6000	503,730.6000	0.3006
		(0.0919)	(242.8736)	(0.1204)	(0.1328)				
	κ-gen	**2.6278**	**13,776.7747**	**0.7432**	–	248,389.3000	496,784.6000	496,805.6000	0.2935
		(0.0220)	**(56.9236)**	**(0.0172)**					
	SM	2.9592	14,091.7755	–	1.4114	248,302.9000	496,611.9000	496,632.9000	0.2911
1991									
		(0.0279)	(204.7070)		(0.0397)	(19.3378)**			
	D	3.5919	13,250.2177	0.7910	–	248,338.6000	496,683.1000	496,704.2000	0.2959
		(0.0385)	(137.6798)	(0.0183)		(90.5883)**			
	GB2	2.4853	14,402.9800	1.2933	1.8549	248,293.3000	496,594.5000	496,622.6000	0.2903
		(0.1043)	(267.4048)	(0.0811)	(0.1349)				
	κ-gen	**2.0851**	**13,400.4939**	**0.5983**	–	246,443.7000	492,893.5000	492,914.5000	0.3374
		(0.0158)	**(60.0387)**	**(0.0142)**					
	SM	2.2627	16,729.0753	–	2.0087	246,487.3000	492,980.5000	493,001.5000	0.3354

Table 2 continued

Wave	Distribution	Parameters[b]				$-\ln L$[c]	AIC	BIC	Gini[d]
		a ($\boldsymbol{\alpha}$)	b ($\boldsymbol{\beta}$)	p ($\boldsymbol{\kappa}$)	q				
1993									
		(0.0196)	(339.3489)		(0.0628)	(127.8631)**			
	D	3.5017	14,725.9254	0.5585	–	246,423.5000	492,853.0000	492,874.0000	0.3375
		(0.0387)	(138.0723)	(0.0112)		(0.3125)			
	GB2	3.4299	14,816.0494	0.5725	1.0340	246,423.3000	492,854.7000	492,882.7000	0.3372
		(0.1323)	(216.3480)	(0.0278)	(0.0620)				

[a] *Boldface entries*: parameter estimates for the κ-generalized distribution
[b] Numbers in *parentheses*: estimated standard errors
[c] Numbers in *parentheses*: likelihood ratio statistics. The critical values of the χ^2 distribution with 1 d.f. are 3.8415 at the 5% level and 6.6349 at the 1% level of significance
[d] Analytic value obtained by substituting the estimated parameters into the relevant expressions. The formulas for the Singh-Maddala, Dagum and GB2 distributions can be found in Kleiber and Kotz (2003)
* Significant at the 5% level
** Significant at the 1% level

Table 3 Estimated distribution functions for the Italian household income in 1995, 1998 and 2000[a]

Wave	Distribution	Parameters[b]				$-\ln L$ [c]	AIC	BIC	Gini [d]
		a ($\boldsymbol{\alpha}$)	b ($\boldsymbol{\beta}$)	p ($\boldsymbol{\kappa}$)	q				
	κ-gen	**2.2448**	**13,232.3586**	**0.6877**	–	240,154.6000	480,315.2000	480,336.2000	0.3307
		(0.0178)	**(59.1460)**	**(0.0153)**					
	SM	2.4898	14,427.1180	–	1.5896	240,167.8000	480,341.5000	480,362.5000	0.3289
1995									
		(0.0223)	(243.9324)		(0.0445)	(29.4995)**			
	D	3.3814	13,689.1872	0.6576	–	240,155.2000	480,316.4000	480,337.4000	0.3310
		(0.0375)	(141.6989)	(0.0142)		(4.4127)*			
	GB2	3.1080	13,944.3106	0.7318	1.1405	240,153.0000	480,314.0000	480,342.0000	0.3297
		(0.1291)	(201.0767)	(0.0408)	(0.0725)				
	κ-gen	**2.1643**	**14,076.5772**	**0.6628**	–	203,917.3000	407,840.7000	407,861.3000	0.3375
		(0.0178)	**(66.3304)**	**(0.0149)**					
1998									
	SM	2.3564	16,262.2511	–	1.7467	204,044.1000	408,094.1000	408,114.7000	0.3360
		(0.0216)	(296.4828)		(0.0506)	(390.7911)**			
	D	3.5656	15,648.1546	0.5542	–	203,896.1000	407,798.2000	407,818.8000	0.3329
		(0.0416)	(147.8028)	(0.0115)		(94.9043)**			
	GB2	5.5177	14,536.7970	0.3335	0.5391	203,848.7000	407,705.3000	407,732.8000	0.3392
		(0.2876)	(141.3052)	(0.0202)	(0.0354)				

Table 3 continued

Wave	Distribution	Parameters[b]				$-\ln L$ [c]	AIC	BIC	Gini [d]
		a ($\boldsymbol{\alpha}$)	b ($\boldsymbol{\beta}$)	p ($\boldsymbol{\kappa}$)	q				
	κ-gen	**2.2549**	**14,492.7122**	**0.6759**	–	223,762.7000	447,531.5000	447,552.4000	0.3272
		(0.0183)	**(65.4196)**	**(0.0154)**					
2000									
	SM	2.4777	16,246.3159	–	1.6713	223,821.1000	447,648.2000	447,669.2000	0.3251
		(0.0226)	(283.2802)		(0.0483)	(138.8074)**			
	D	3.5313	15,465.6011	0.6095	–	223,755.5000	447,517.0000	447,537.9000	0.3254
		(0.0399)	(150.3455)	(0.0129)		(7.5727)**			
	GB2	3.9402	15,110.3582	0.5337	0.8462	223,751.7000	447,511.4000	447,539.4000	0.3272
		(0.1660)	(181.8476)	(0.0283)	(0.0510)				

[a] *Boldface entries*: parameter estimates for the κ-generalized distribution
[b] Numbers in *parentheses*: estimated standard errors
[c] Numbers in *parentheses*: likelihood ratio statistics. The critical values of the χ^2 distribution with 1 d.f. are 3.8415 at the 5% level and 6.6349 at the 1% level of significance
[c] Analytic value obtained by substituting the estimated parameters into the relevant expressions. The formulas for the Singh-Maddala, Dagum and GB2 distributions can be found in Kleiber and Kotz (2003)
* Significant at the 5% level
** Significant at the 1% level

Table 4 Estimated distribution functions for the Italian household income in 2002, 2004 and 2006[a]

Wave	Distribution	Parameters[b]				$-\ln L$[c]	AIC	BIC	Gini[d]
		a ($\boldsymbol{\alpha}$)	b ($\boldsymbol{\beta}$)	p ($\boldsymbol{\kappa}$)	q				
2002									
	κ-gen	**2.2943**	**14,630.8555**	**0.6895**	–	222,026.3000	444,058.7000	444,079.6000	0.3243
		(0.0189)	**(65.9872)**	**(0.0157)**					
	SM	2.5107	16,397.4011	–	1.6659	222,085.2000	444,176.4000	444,197.4000	0.3215
		(0.0233)	(285.9187)		(0.0487)	(162.1081)**			
	D	3.5646	15,595.3330	0.6131	–	222,010.8000	444,027.6000	444,048.6000	0.3220
		(0.0396)	(148.2517)	(0.0128)		(13.3356)**			
	GB2	4.1261	15,111.0365	0.5146	0.7997	222,004.2000	444,016.3000	444,044.2000	0.3245
		(0.1767)	(177.2646)	(0.0273)	(0.0488)				
2004									
	κ-gen	**2.3385**	**14,926.5593**	**0.7354**	–	213,811.2000	427,628.5000	427,649.5000	0.3267
		(0.0197)	**(69.2921)**	**(0.0165)**					
	SM	2.6368	15,196.7667	–	1.3977	213,842.6000	427,691.1000	427,712.1000	0.3260
		(0.0248)	(237.1617)		(0.0378)	(56.6202)**			
	D	3.3847	15,196.7547	0.6911	–	213,814.6000	427,635.2000	427,656.2000	0.3258
		(0.0392)	(166.4068)	(0.0159)		(0.7086)			
	GB2	3.4960	15,196.7467	0.6604	0.9581	213,814.2000	427,636.5000	427,664.4000	0.3259
		(0.1459)	(193.9699)	(0.0367)	(0.0579)				
2006									
	κ-gen	**2.3674**	**15,761.5688**	**0.7168**	–	206,278.6000	412,563.1000	412,584.0000	0.3196
		(0.0205)	**(73.6083)**	**(0.0169)**					

Table 4 continued

Wave	Distribution	Parameters[b]				$-\ln L$[c]	AIC	BIC	Gini[d]
		a ($\boldsymbol{\alpha}$)	b ($\boldsymbol{\beta}$)	p ($\boldsymbol{\kappa}$)	q				
	SM	2.6349	16,703.2043	–	1.5073	206,312.5000	412,631.0000	412,651.9000	0.3175
		(0.0254)	(276.8047)		(0.0435)	(77.6099)**			
	D	3.5163	16,298.6836	0.6625	–	206,274.5000	412,555.1000	412,575.9000	0.3183
		(0.0411)	(171.0779)	(0.0151)		(1.6888)			
	GB2	3.7006	16,133.1813	0.6219	0.9243	206,273.7000	412,555.4000	412,583.2000	0.3192
		(0.1532)	(204.9994)	(0.0336)	(0.0559)				

[a] *Boldface entries*: parameter estimates for the κ-generalized distribution
[b] Numbers in *parentheses*: estimated standard errors
[c] Numbers in *parentheses*: likelihood ratio statistics. The critical values of the χ^2 distribution with 1 d.f. are 3.8415 at the 5% level and 6.6349 at the 1% level of significance
[d] Analytic value obtained by substituting the estimated parameters into the relevant expressions. The formulas for the Singh-Maddala, Dagum and GB2 distributions can be found in Kleiber and Kotz (2003)
* Significant at the 5% level
** Significant at the 1% level

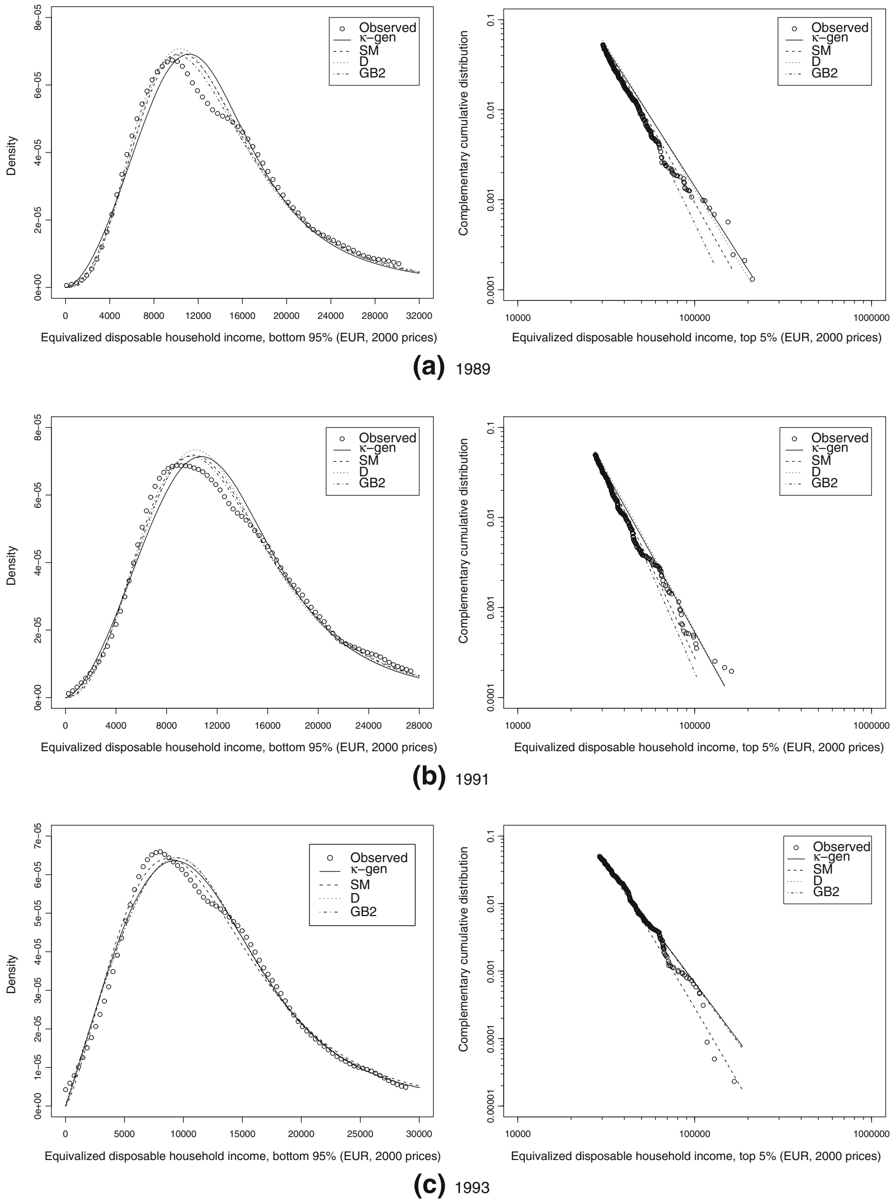


Fig. 4 Observed and predicted density (*left*) and log–log complementary distribution (*right*) for the Italian household income data in 1989, 1991 and 1993. The density has been estimated using an Epanechnikov kernel

κ values holding α constant are associated with increased inequality, whereas higher values of α holding κ constant leads to declines in inequality.[11]

[11] The region around the origin of the κ-generalized distribution is governed by α, the upper tail by both α and κ. In particular, as shown in Fig. 1, increasing κ leads to a thicker upper tail, whereas increasing α lowers

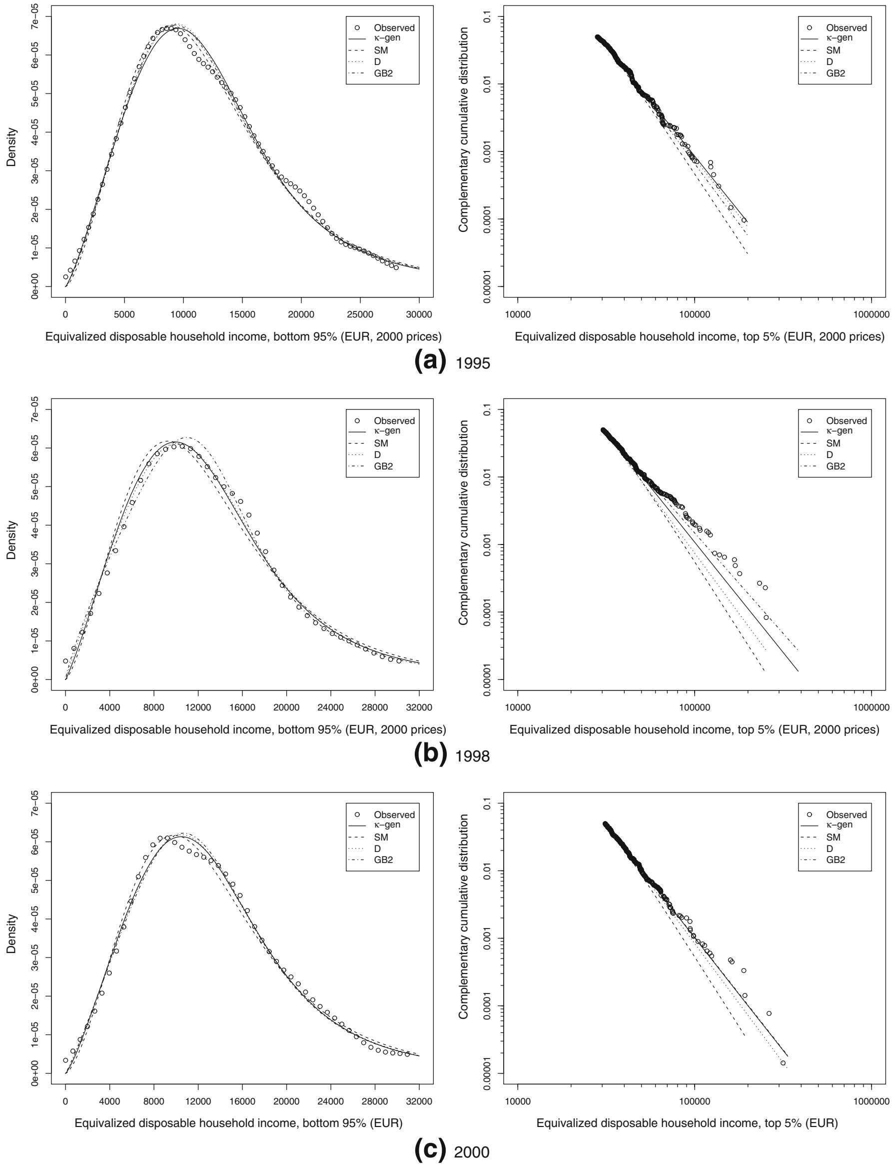


Fig. 5 Observed and predicted density (*left*) and log–log complementary distribution (*right*) for the Italian household income data in 1995, 1998 and 2000. The density has been estimated using an Epanechnikov kernel

Footnote 11 continued
both the tails and yields a greater concentration of probability mass around the peak of the distribution. Hence κ is an "inequality" parameter, and α can be called "equality" parameter because the Gini coefficients increases with the former and decreases with the latter. The same results (available upon request) hold true for the other measures of inequality derived in Sect. 2.3.

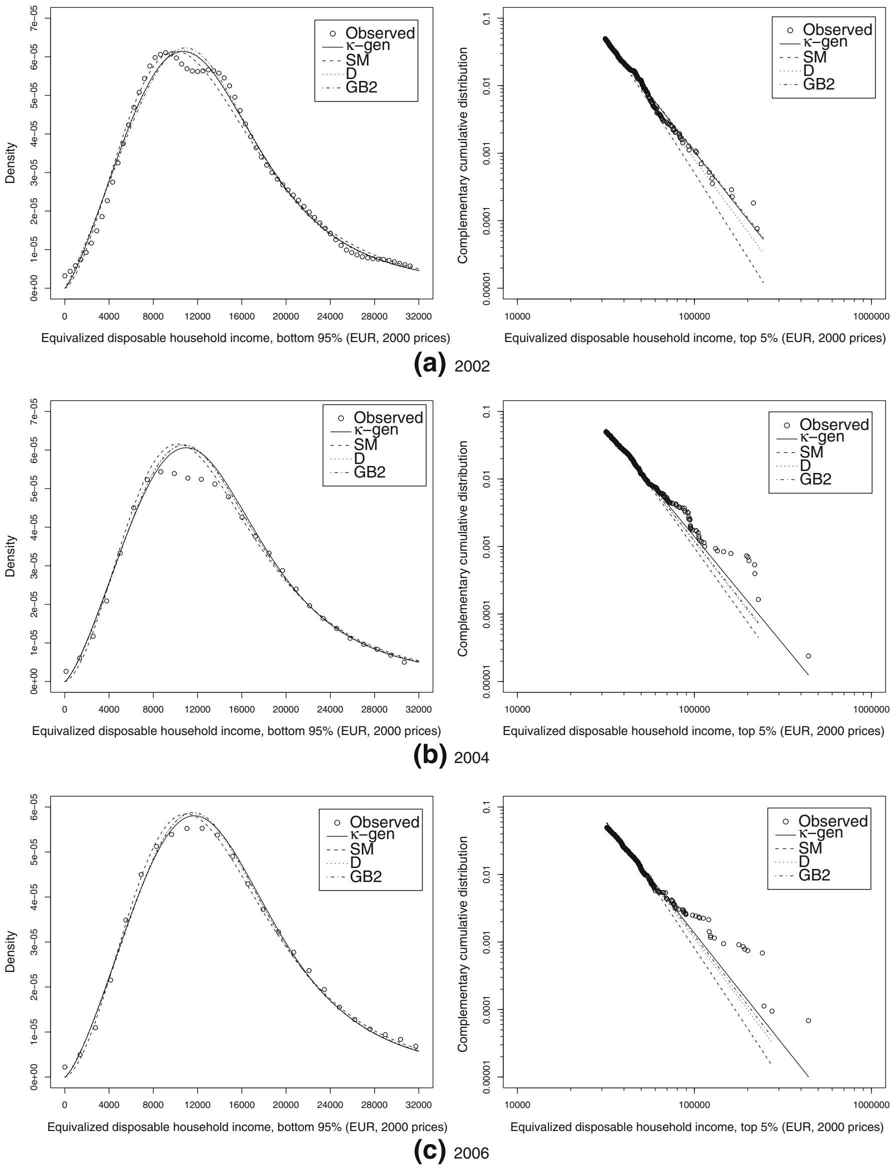


Fig. 6 Observed and predicted density (*left*) and log–log complementary distribution (*right*) for the Italian household income data in 2002, 2004 and 2006. The density has been estimated using an Epanechnikov kernel

For comparison, the results of fitting other existing functional forms that have been considered successful in describing the income size distribution are also shown in the tables and figures above. Namely, these models are the four-parameter generalized beta II distribution (GB2) introduced by McDonald (1984), which has the density

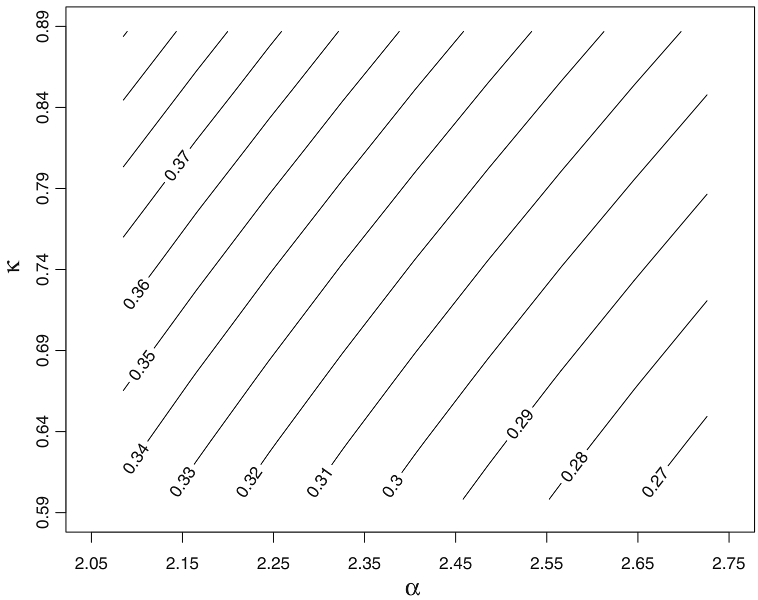


Fig. 7 Gini coefficient as a function of the κ-generalized parameters

$$f(x; a, b, p, q) = \frac{ax^{ap-1}}{b^{ap} B(p, q) \left[1 + (x/b)^a\right]^{p+q}}, \quad x \geq 0, \quad a, b, p, q > 0,$$

where $B(\cdot, \cdot)$ is the beta function, and the three-parameter Singh-Maddala (SM) and Dagum type I (D) distributions, corresponding to the special cases

$$\text{SM}(a, b, q) = \text{GB2}(a, b, 1, q) \quad \text{and} \quad \text{D}(a, b, p) = \text{GB2}(a, b, p, 1).$$ [12]

By inspection of AIC and BIC values, it emerges that among the three-parameter models the κ-generalized provides the best fit in 1995 and 2004, while the Singh-Maddala is the favorite in 1989 and 1991. The Dagum distribution results in a better fit in 5 out of nine cases, i.e. in 1993, 1998, 2000, 2002 and 2006. Using a likelihood ratio test at the 5% level of significance,[13] the four-parameter GB2 provides a statistically significant improvement over its nested Singh-Maddala distribution in all cases, whereas the differences between the GB2 and Dagum are not

[12] These relationships can be seen in greater detail in McDonald and Xu (1995).

[13] The likelihood ratio provides the basis for comparing nested models. The asymptotic distribution of $2\left[\ln L\left(\hat{\theta}_{\mathrm{U}}\right) - \ln L\left(\hat{\theta}_{\mathrm{R}}\right)\right]$ is χ^2 with degrees of freedom (d.f.) equal to the number of independent restrictions imposed on the more general model in order to yield the nested one. $\hat{\theta}_{\mathrm{U}}$ and $\hat{\theta}_{\mathrm{R}}$, respectively, denote maximum likelihood estimators of the unrestricted and restricted model. This test can not be used to compare nonnested models. In the tables, asterisks are placed next to the likelihood ratio values if the improvement gained in adding a further parameter is of practical significance.

significant in 1993, 2004 and 2006, and also in 1995 if one lowers the significance level at the 1%.

To examine if the differences between our fitted model and the others are statistically significant, the Vuong (1989) approach to model selection in the case of nonnested hypotheses has been implemented. The approach is probabilistic and is based on testing the null hypothesis that the competing models are equally close to the true data generating process against the alternative hypothesis that one model is closer. The resulting likelihood ratio statistic is asymptotically distributed as a standard normal under the null hypothesis; the actual test is then

$$\text{under } H_0 : V = \frac{\mathcal{R}}{\hat{\sigma}\sqrt{n}} \xrightarrow{d} N(0, 1), \tag{17}$$

where

$$\mathcal{R} = \sum_{i=1}^{n} \ln \frac{f_1\left(x_i; \hat{\theta}_1\right)}{f_2\left(x_i; \hat{\theta}_2\right)},$$

being $\hat{\theta}_j$, $j = 1, 2$, the maximum likelihood estimator of the unknown parameters for the model F_j, and

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\ln \frac{f_1\left(x_i; \hat{\theta}_1\right)}{f_2\left(x_i; \hat{\theta}_2\right)}\right]^2 - \left[\frac{1}{n}\sum_{i=1}^{n}\ln \frac{f_1\left(x_i; \hat{\theta}_1\right)}{f_2\left(x_i; \hat{\theta}_2\right)}\right]^2}.$$

Chosen a critical value z from the standard normal distribution corresponding to the desired level of significance, if $|V| \leq z$, the null that models are the same and cannot be rejected, whereas if $V > z$, the model F_1 can be considered better than model F_2, and the reverse is true if $V < -z$. Since the Vuong test is sensitive to the number of estimated parameters in each model, we have chosen to apply the correction for the model dimensionality

$$\tilde{\mathcal{R}} \equiv \mathcal{R} - \left[\left(\frac{d_1}{2}\right)\ln n - \left(\frac{d_2}{2}\right)\ln n\right], \tag{18}$$

where d_j is the number of estimated parameters within distribution j, which corresponds to Schwarz information criterion (Vuong 1989, p. 318). The adjusted statistic (17) has the same asymptotic properties of (18).

Table 5 reports the calculated Vuong statistics for comparison between our model and the other distributions. In nearly all of the cases considered (seven out of nine, from 1993 to 2006) the κ-generalized distribution is observationally equivalent to the Singh-Maddala if one takes 5% as the relevant significance level; furthermore, in three out of these seven cases (1998, 2000 and 2002) its superiority as a descriptive model is found to be statistically significant. In the same number of cases similar results are obtained when comparing to the Dagum distribution, whereas the four-parameter GB2

Table 5 Vuong test for nonnested model selection[a]

Distributions	Wave								
	1989	1991	1993	1995	1998	2000	2002	2004	2006
κ-gen–SM	−5.3596**	−3.3274**	1.5329	0.6325	5.3735**	2.5112*	2.1410*	1.2371	1.2072
κ-gen–D	−3.9172**	−2.2152*	−1.3121	0.0928	−1.2671	−1.0334	−1.4967	0.4724	−0.5927
κ-gen–GB2	−3.4640**	−2.4435*	−1.1135	0.3299	−2.1159*	−0.6268	−1.0781	1.3432	−0.0683

[a] The critical values of the Vuong statistic are 1.9600 at the 5% level, and 2.5758 at the 1% level of significance
* Significant at the 5% level
** Significant at the 1% level

distribution provides a statistically significant better fit relative to the κ-generalized only in 1998. There are just two exceptions to this rule. Indeed, in 1989 and 1991, our model is significantly outperformed by every other distribution, with the GB2 being selected as the one providing the best fit. The figures of Table 1 show that between 1989 and 1991 the middle mass of the income density increased, so reducing the thickness of the tails and therefore the income inequality. Consequently, the GB2 density with its additional parameter is more appropriate than the best three-parameter model to reflect the impact of economic fluctuations on the income distribution in those years. However, if one lowers the significance level at 1% in almost every situation (eight cases out of nine, from 1991 to 2006) the performance of our distribution is at least as good as that of the others (except for the 1991 wave, where the Singh-Maddala distribution performs significantly better).

Finally, to get a more formal idea of the fit of each distributional model, we look at the informational entropy of the observed data and compare it to the entropy of different fitted distribution functions.[14] The results are listed in Table 6 and visually displayed in Fig. 8. What is immediately notable is that the entropy for the Singh-Maddala, Dagum and GB2 distributions is much larger than the entropy of the data. This fact means that there are relevant moment constraints that have been neglected in these distributional models, i.e. there is testable information relevant to the description of the data that is not being considered. By contrast, the κ-generalized distribution is found to give a better entropy match, even though the entropy based on this model is now slightly lower than the empirical one. This could suggest that this model is mildly over-constrained, but at least no relevant information is being left unused.

4 Summary

One of the main objectives of research on income distribution is to provide a mathematical description of the size distribution of income for approximating the underlying "true" distribution. Starting from the Pareto contribution, a wide variety of functional forms have been considered as possible models for the distribution of personal income by size, and other approaches can no doubt be suggested and deserve attention.

In this work, we have proposed a three-parameter distribution based on the κ-deformation of the standard exponential and logarithm functions that follows from Kaniadakis (2001, 2002, 2005). Expressions for the moments and various tools for the measurement of income inequality are given that are functions of the parameters in the model. Given estimates of the parameters, these expressions can be used to estimate corresponding population characteristics of interest as well as to provide indirect checks on the validity of the parameter estimation. The model is able to describe the

[14] Comparing the entropy of the data to the entropy based on a particular probability distribution can provide relevant insights into the appropriateness of the distributional model given the data. If the entropy of the data is lower than the entropy based on an assumed distribution function, then there is information in the data that is not being taken advantage of to specify the distribution function. On the other hand, if the entropy of the data is greater than that based on the hypothetical distribution function, more structure is being assumed than is justified by the data. See Jaynes (1978) for a more thorough discussion of this approach.

Table 6 Entropy of the observed data and entropy based on each distributional model using parameter estimates, 1989–2006[a]

	Wave								
	1989	1991	1993	1995	1998	2000	2002	2004	2006
Data	8.4491	8.3876	8.4448	8.5379	8.4190	8.4666	8.4282	8.4610	8.4249
	Predicted								
κ-gen[b]	8.2700	8.1884	8.2718	8.2485	8.3262	8.3296	8.3324	8.3614	8.3958
SM[c]	10.2618	10.1816	10.2698	10.2457	10.3275	10.3279	10.3305	10.3587	10.3927
D[d]	10.2672	10.1893	10.2728	10.2494	10.3201	10.3281	10.3302	10.3603	10.3944
GB2[e]	10.2555	10.1791	10.2723	10.2474	10.3277	10.3305	10.3334	10.3613	10.3957

[a] The empirical entropy has been obtained by replacing the discrete probabilities f_i, $i = 1, \ldots, m$, in the Shannon entropy measure $S = -\sum_{i=1}^{m} f_i \ln f_i$ by maximum likelihood (ML) estimates $\hat{f}_i = \frac{n_i}{n} = \frac{n_i}{\sum_{i=1}^{m} n_i}$, $n \geq m$, where n_i is the frequency of realization i in the sample of n observations

[b] Functional form: $S = -\ln \frac{\alpha}{\beta^{\alpha}} - (\alpha - 1)\left[\ln \beta - \frac{\gamma + \psi\left(\frac{1}{2\kappa}\right) + \ln(2\kappa) + \kappa}{\alpha}\right] - \frac{1}{2}\psi\left(1 + \frac{1}{4\kappa}\right) + \frac{1}{2}\psi\left(\frac{1}{2} + \frac{1}{4\kappa}\right) + \kappa - 1$

[c] Functional form: $S = -\ln(aq) - (a-1)\frac{a \ln b - \psi(q) - \gamma}{a} + \frac{1}{q} + 1$

[d] Functional form: $S = -\ln(ap) - (ap-1)\frac{a \ln b + \psi(p) + \gamma}{a} + (p+1)\left[\psi(p+1) + \gamma\right]$

[e] Functional form: $S = -\ln(a) - (ap-1)\frac{a \ln b + \psi(p) - \psi(q)}{a} + \ln B(p,q) + (p+q)\left[\psi(p+q) - \psi(q)\right]$

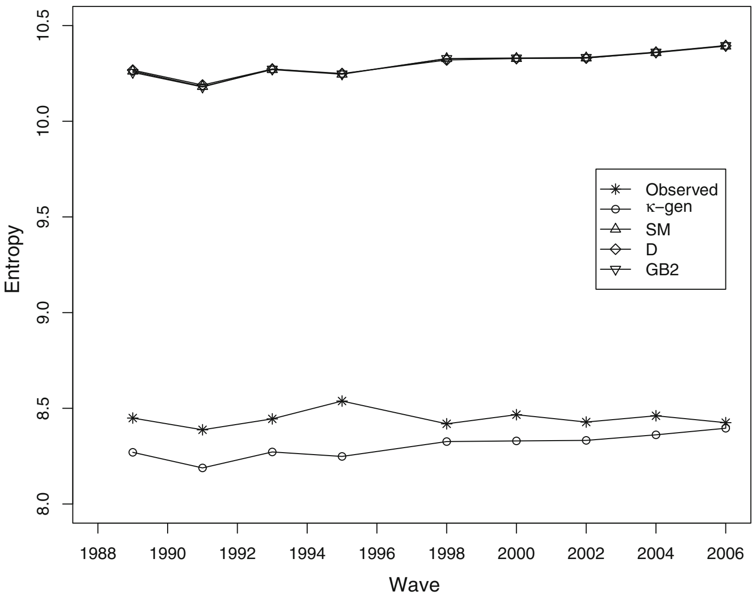


Fig. 8 Observed and predicted entropy across the years

entire income range, including the Pareto upper tail, and fits the Italian income data from 1989 to 2006 fairly well. We find that in a satisfactory number of cases the performance of this model is not to be considered inferior from the statistical point of view to that of the Singh-Maddala, Dagum and generalized beta II distributions, whereas its ability of matching the entropy of the data dominates these others.

It is important to note that income in other countries is likely to be distributed differently. Furthermore, using different definitions of income can cause relevant changes in the observed distribution (e.g. accounting for the government redistribution generally alters the distribution of pre-tax household income in many countries). Therefore, further analysis using both data from a diverse set of countries and different definitions of income are needed in order to draw definitive conclusions about our and the other distributional models. This will be the next step in our research agenda.

Acknowledgements The authors would like to thank two anonymous referees, whose insightful comments helped to improve this article greatly.

Appendix A: Maximum entropy derivation of the κ-generalized distribution

The Lagrangian of the continuous maximum entropy program (2–4) is given by

$$\mathcal{L} = S - \sum_{j=0}^{3} \lambda_j \left[\int_0^\infty \phi_j \left(x; \boldsymbol{\theta} \right) f \left(x \right) - C_j \right],$$

where λ_j are multipliers. The expressions for the functions $\phi_j(x;\boldsymbol{\theta})$, involving the parameter vector $\boldsymbol{\theta}$, and the constants C_j can be recovered by matching terms in the set of constraints (3) and (4). In particular, the normalization constraint correspond to $j = 0$ by setting $\phi_0(x;\boldsymbol{\theta})$ and C_0 equal to 1.

The first-order condition

$$\frac{\partial \mathcal{L}}{\partial f(x)} = -\ln f(x) - (1+\lambda_0) - \sum_{j=1}^{3} \lambda_j \phi_j(x;\boldsymbol{\theta}) = 0$$

implies that the solution has the functional form

$$f(x;\boldsymbol{\theta}) = \frac{1}{\Omega(\boldsymbol{\theta})} \exp\left[-\sum_{j=1}^{3} \lambda_j \phi_j(x;\boldsymbol{\theta})\right], \tag{A.1}$$

where

$$\Omega(\boldsymbol{\theta}) = \exp(1+\lambda_0) = \int_0^{\infty} \exp\left[-\sum_{j=1}^{3} \lambda_j \phi_j(x;\boldsymbol{\theta})\right] \tag{A.2}$$

is the partition function that normalizes the density and can be determined from Eq. 3.

In order to determine the values of the multipliers as a function of the parameters $\boldsymbol{\theta} = \{\alpha, \beta, \kappa\}$, we observe that for each j it holds (Kaniadakis, 2009)

$$\lambda_j = \frac{\partial S}{\partial C_j}. \tag{A.3}$$

Hence, taking log of Eq. 5, multiplying it by $[-f(x)]$ and integrating between 0 and ∞, we get

$$\begin{aligned} S &= \ln \frac{\beta^{\alpha}}{\alpha} - (\alpha - 1) \int_0^{\infty} \ln x f(x)\,\mathrm{d}x + \frac{1}{2} \int_0^{\infty} \ln\left[1 + \kappa^2 \left(\frac{x}{\beta}\right)^{2\alpha}\right] f(x)\,\mathrm{d}x \\ &\quad - \frac{1}{\kappa} \int_0^{\infty} \sinh^{-1}\left[-\kappa \left(\frac{x}{\beta}\right)^{\alpha}\right] f(x)\,\mathrm{d}x \\ &= \ln \frac{\beta^{\alpha}}{\alpha} - (\alpha - 1) C_1 + \frac{1}{2} C_2 - \frac{1}{\kappa} C_3, \end{aligned}$$

which, applying (A.3), produces

$$\lambda_1 = 1 - \alpha, \quad \lambda_2 = \frac{1}{2} \quad \text{and} \quad \lambda_3 = -\frac{1}{\kappa}. \tag{A.4}$$

Finally, by inserting these results into (A.2) and solving the integral yields $\Omega(\boldsymbol{\theta}) = \frac{\beta^\alpha}{\alpha}$, which can be substituted for together with (A.4) into (A.1) leaving us with Eq. 5 in the main text.

Appendix B: Proof of Theorem 1

Since the Lorenz curve is invariant under scale changes, the parameter β does not enter expression (10) and can be chosen as 1 without loss of generality. Furthermore, we state the following:

Lemma 1 **(Hardy et al. 1929; see also Marshall and Olkin 1979)** *Let X_i, $i = 1, 2$, be positive random variables having finite mean m_{X_i}. Then $X_1 \leq_L X_2$ if and only if*

$$\mathbb{E}\left[\psi\left(\frac{X_1}{m_{X_1}}\right)\right] \leq \mathbb{E}\left[\psi\left(\frac{X_2}{m_{X_2}}\right)\right]$$

for all continuous and convex functions $\psi : \mathbb{R}_+ \to \mathbb{R}$ for which the expectation exists.

Now we can prove the first direction of Theorem 1, i.e. necessity of conditions $\alpha_1 \geq \alpha_2$ and $\frac{\alpha_1}{\kappa_1} \geq \frac{\alpha_2}{\kappa_2}$ for Lorenz-ordering of κ-generalized distributions. Consider the family of continuous convex functions

$$\Psi(x) = \frac{x^{t+1} - 1}{t(t+1)}, \quad x > 0, \quad -\infty < t < \infty, \quad t \neq -1, 0.$$

As can be seen in Taillie (1981), corresponding to Ψ one can obtain the family of inequality measures

$$\begin{aligned} H_t(X) &= \mathbb{E}\left[\Psi\left(\frac{X}{m_X}\right)\right] \\ &= \frac{1}{t(t+1)}\left[\frac{\mathbb{E}\left(X^{t+1}\right)}{m_X^{t+1}} - 1\right], \quad x > 0, \quad -\infty < t < \infty, \quad t \neq -1, 0 \end{aligned}$$

that preserve the Lorenz-ordering. This family includes some standard indices as special cases, such as one-half the squared coefficient of variation (H_1) and the Theil index (H_0). From Lemma 1, we know that

$$H_t(X_1) \leq H_t(X_2), \tag{B.1}$$

where

$$
\begin{aligned}
&H_t\left(X_1\right)\\
&=\frac{1}{t\left(t+1\right)}\left\{\frac{\left(2\kappa_1\right)^{-\frac{t+1}{\alpha_1}}\frac{\Gamma\left(1+\frac{t+1}{\alpha_1}\right)}{1+\frac{t+1}{\alpha_1}\kappa_1}\frac{\Gamma\left(\frac{1}{2\kappa_1}-\frac{t+1}{2\alpha_1}\right)}{\Gamma\left(\frac{1}{2\kappa_1}+\frac{t+1}{2\alpha_1}\right)}}{\left[\left(2\kappa_1\right)^{-\frac{1}{\alpha_1}}\frac{\Gamma\left(1+\frac{1}{\alpha_1}\right)}{1+\frac{\kappa_1}{\alpha_1}}\frac{\Gamma\left(\frac{1}{2\kappa_1}-\frac{1}{2\alpha_1}\right)}{\Gamma\left(\frac{1}{2\kappa_1}+\frac{1}{2\alpha_1}\right)}\right]^{t+1}}-1\right\}\\
&=\frac{1}{t\left(t+1\right)}\left\{\frac{\Gamma\left(1+\frac{t+1}{\alpha_1}\right)}{1+\frac{t+1}{\alpha_1}\kappa_1}\frac{\Gamma\left(\frac{1}{2\kappa_1}-\frac{t+1}{2\alpha_1}\right)}{\Gamma\left(\frac{1}{2\kappa_1}+\frac{t+1}{2\alpha_1}\right)}\left[\frac{1+\frac{\kappa_1}{\alpha_1}}{\Gamma\left(1+\frac{1}{\alpha_1}\right)}\frac{\Gamma\left(\frac{1}{2\kappa_1}+\frac{1}{2\alpha_1}\right)}{\Gamma\left(\frac{1}{2\kappa_1}-\frac{1}{2\alpha_1}\right)}\right]^{t+1}-1\right\}
\end{aligned}
$$

follows from (9) and is not defined outside the interval $\left(-\alpha_1-1,\frac{\alpha_1}{\kappa_1}-1\right)$. A similar expression obtains for $H_t\left(X_2\right)$. Recalling that $\lim_{z\to 0}\Gamma\left(z\right)=+\infty$, the terms $\Gamma\left(1+\frac{t+1}{\alpha_1}\right)$ and $\Gamma\left(\frac{1}{2\kappa_1}-\frac{t+1}{2\alpha_1}\right)$ tend to infinity as t approaches $-\alpha_1-1$ from above and $\frac{\alpha_1}{\kappa_1}-1$ from below, respectively. Similarly, $H_t\left(X_2\right)$ is not defined outside the interval $\left(-\alpha_2-1,\frac{\alpha_2}{\kappa_2}-1\right)$, and tends to infinity as $t\to-\alpha_2-1$ and $t\to\frac{\alpha_2}{\kappa_2}-1$. Hence, in conjunction with (B.1), it follows that $-\alpha_1-1\le-\alpha_2-1$ and $\frac{\alpha_1}{\kappa_1}-1\ge\frac{\alpha_2}{\kappa_2}-1$, or equivalently $\alpha_1\ge\alpha_2$ and $\frac{\alpha_1}{\kappa_1}\ge\frac{\alpha_2}{\kappa_2}$.

It remains to prove the reverse direction, i.e. that the conditions of Theorem 1 are also sufficient. Since the quantile function is available in a simple closed form, we can check for star-shaped ordering $X_1\le_* X_2$ (e.g. Arnold 1987) by verifying that the derivative with respect to u of the ratio

$$
\frac{F_{X_2}^{-1}\left(u;\alpha_2,\kappa_2\right)}{F_{X_1}^{-1}\left(u;\alpha_1,\kappa_1\right)}=\frac{\left[\ln_{\kappa_2}\left(\frac{1}{1-u}\right)\right]^{\frac{1}{\alpha_2}}}{\left[\ln_{\kappa_1}\left(\frac{1}{1-u}\right)\right]^{\frac{1}{\alpha_1}}}=\frac{\left[\frac{(1-u)^{-\kappa_2}-(1-u)^{\kappa_2}}{2\kappa_2}\right]^{\frac{1}{\alpha_2}}}{\left[\frac{(1-u)^{-\kappa_1}-(1-u)^{\kappa_1}}{2\kappa_1}\right]^{\frac{1}{\alpha_1}}}
$$

is nonnegative. After some straightforward manipulations, one ends up with the result

$$
\frac{\alpha_1}{\kappa_1}\left[\frac{(1-u)^{-\kappa_1}-(1-u)^{\kappa_1}}{(1-u)^{-\kappa_1}+(1-u)^{\kappa_1}}\right]\ge\frac{\alpha_2}{\kappa_2}\left[\frac{(1-u)^{-\kappa_2}-(1-u)^{\kappa_2}}{(1-u)^{-\kappa_2}+(1-u)^{\kappa_2}}\right],
$$

which holds true assuming that $\alpha_1\ge\alpha_2$ and $\frac{\alpha_1}{\kappa_1}\ge\frac{\alpha_2}{\kappa_2}$. Therefore, conditions (11) are sufficient for the star-shaped order of X_1 with respect to X_2, which in turn implies the Lorenz-ordering $X_1\le_L X_2$ (e.g. Chandra and Singpurawalla 1981).

Figure 9 shows the decomposition of (α_1,$\frac{\alpha_1}{\kappa_1}$)-space into regions where X_1 dominates or is dominated by X_2, given α_2 and $\frac{\alpha_2}{\kappa_2}$, and where instead the two distributions are not comparable because their Lorenz curves cross.

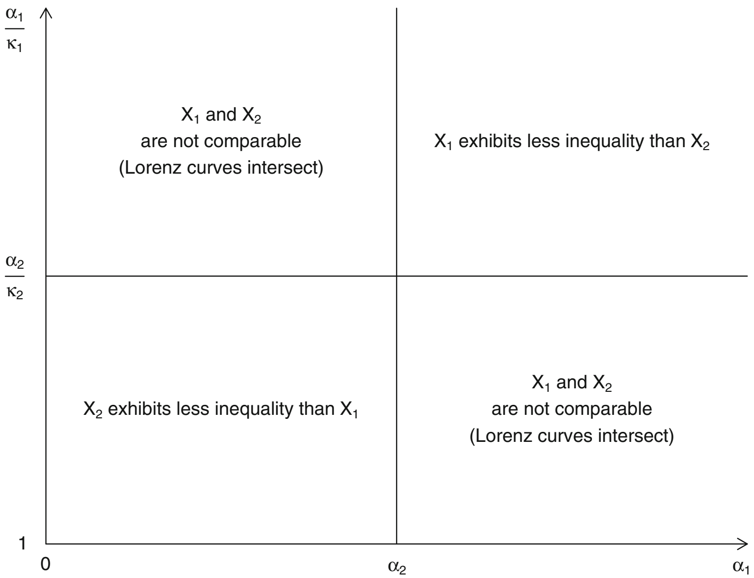


Fig. 9 Lorenz-ordering within the κ-generalized family

References

Aitchison J, Brown JAC (1954) On criteria for descriptions of income distribution. Metroeconomica 6: 81–144

Aitchison J, Brown JAC (1957) The lognormal distribution with special reference to its use in economics. Cambridge University Press, New York

Akaike H (1973) Information theory and an extension of the likelihood ratio principle. In: Petrov BN, Csaki F (eds) Proceedings of the second international symposium of information theory. Akademiai Kiado, Budapest, pp 257–281

Amiel Y, Cowell FA, Polovin A (1996) Inequality among the kibbutzim. Economica 63: S63–S85

Arnold BC (1983) Pareto distributions. International Co-operative Publishing House, Fairland

Arnold BC (1987) Majorization and the Lorenz order: a brief introduction. Springer-Verlag, Berlin

Arnold BC, Laguna L (1977) On generalized Pareto distributions with applications to income data. Iowa State University Press, Ames

Atkinson AB (1970) On the measurement of inequality. J Econ Theory 2:244–263

Atoda N, Suruga T, Tachibanaki T (1988) Statistical inference of functional forms for income distribution. Econ Stud Q 39:14–40

Bartels CPA (1977) Economic aspects of regional welfare: income distribution and unemployment. Martinus Nijhoff, Leiden

Bordley RF, McDonald JB, Mantrala A (1996) Something new, something old: parametric models for the size distribution of income. J Income Distrib 6:91–103

Brachmann K, Stich A, Trede M (1996) Evaluating parametric income distribution models. Alleg Stat Arch 80:285–298

Brandolini A (1999) The distribution of personal income in post-war Italy: source description, data quality, and the time pattern of income inequality. Giorn Econ 58:183–239

Champernowne DG (1953) A model of income distribution. Econ J 63:318–351

Chandra M, Singpurwalla ND (1981) Relationships between some notions which are common to reliability theory and economics. Math Oper Res 6:113–121

Cowell FA (1980a) Generalized entropy and the measurement of distributional change. Eur Econ Rev 13:147–159

Cowell FA (1980b) On the structure of additive inequality measures. Rev Econ Stud 47:521–531

Cowell FA (1995) Measuring inequality. Prentice Hall/Harvester Wheatsheaf, Hemel Hempstead

Cowell FA, Kuga K (1981a) Additivity and the entropy concept: an axiomatic approach to inequality measurement. J Econ Theory 25:131–143

Cowell FA, Kuga K (1981b) Inequality measurement: an axiomatic approach. Eur Econ Rev 15:287–305

Dagum C (1977) A new model of personal income distribution: specification and estimation. Econ Appl 30:413–436

Espinguet P, Terraza M (1983) Essai d'extrapolation des distributions de salaires français. Econ Appl 36:535–561

Esteban JM (1986) Income-share elasticity and the size distribution of income. Int Econ Rev 27:439–444

Gastwirth JL (1971) A general definition of the Lorenz curve. Econometrica 39:1037–1039

Gibrat R (1931) Les inégalités économiques. Applications: aux inégalités des richesses, à la concentration des entreprises, aux population des villes, aux statistiques des familles, etc., d'une loi nouvelle: la loi de l'effet proportionnel. Librairie du Recueil Sirey, Paris

Gini C (1914) Sulla misura della concentrazione e della variabilità dei caratteri. Atti del Reale Istituto veneto di scienze, lettere ed arti 73:1201–1248 (trans: On the measurement of concentration and variability of characters. Metron 63:3–38, 2005)

Hardy GH, Littlewood JE, Pólya G (1929) Some simple inequalities satisfied by convex functions. Messenger Math 58:145–152

Jaynes ET (1957a) Information theory and statistical mechanics. Phys Rev 106:620–630

Jaynes ET (1957b) Information theory and statistical mechanics. II. Phys Rev 108:171–190

Jaynes ET (1978) Where do we stand on maximum entropy? In: Levine RD, Tribus M (eds) The maximum entropy formalism. MIT Press, Cambridge pp 18–115

Kakwani N (1980) Income inequality and poverty: methods of estimation and policy applications. Oxford University Press, New York

Kaniadakis G (2001) Non-linear kinetics underlying generalized statistics. Phys A 296:405–425

Kaniadakis G (2002) Statistical mechanics in the context of special relativity. Phys Rev E 66:056125

Kaniadakis G (2005) Statistical mechanics in the context of special relativity. II. Phys Rev E 72:036108

Kaniadakis G (2009) Maximum entropy principle and power-law tailed distributions. Eur Phys J B 70:3–13

Kapur JN (1989) Maximum-entropy models in science and engineering. Wiley, New York

Kleiber C (1996) Dagum vs. Singh-Maddala income distributions. Econ Lett 53:265–268

Kleiber C, Kotz S (2003) Statistical size distributions in economics and actuarial sciences. Wiley, New York

Leipnik RB (1990) A maximum relative entropy principle for distribution of personal income with derivations of several known income distributions. Commun Stat Theory 19:1003–1036

Lorenz MO (1905) Methods of measuring the concentration of wealth. Pub Am Stat Assn 9:209–219

Majumder A, Chakravarty SR (1990) Distribution of personal income: development of a new model and its application to U.S. income data. J Appl Econ 5:189–196

Mandelbrot B (1960) The Pareto-Lévy law and the distribution of income. Int Econ Rev 1:79–106

Marshall AW, Olkin I (1979) Inequalities: theory of majorization and its applications. Academic Press, New York

McDonald JB (1984) Some generalized functions for the size distribution of income. Econometrica 52:647–665

McDonald JB, Ransom MR (1979) Functional forms, estimation techniques and the distribution of income. Econometrica 47:1513–1525

McDonald JB, Xu YJ (1995) A generalization of the beta distribution with applications. J Econom 66:133–152 (Errata: J Econom 69:427–428)

Metcalf CE (1972) An econometric model of the income distribution. Markham Publishing Company, Chicago

Ord JK, Patil GP, Taillie C (1981) The choice of a distribution to describe personal incomes. In: Taillie C, Patil GP, Baldessari BA (eds) Statistical distributions in scientific work, vol 6. D. Reidel Publishing Company, Dordrecht pp 193–201

R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org

Rajaonarison D (2008) Deterministic heterogeneity in tastes and product differentiation in the K-logit model. Econ Lett 100:396–398

Rajaonarison D, Bolduc D, Jayet H (2005) The K-deformed multinomial logit model. Econ Lett 86:13–20

Salem ABZ, Mount TD (1974) A convenient descriptive model of income distribution: the gamma density. Econometrica 42:1115–1127

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464
Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423, 623–657
Shorrocks AF (1980) The class of additively decomposable inequality measures. Econometrica 48:613–625
Singh SK, Maddala GS (1976) A function for size distribution of incomes. Econometrica 44:963–970
Tachibanaki T, Suruga T, Atoda N (1997) Estimations of income distribution parameters for individual observations by maximum likelihood method. J Jpn Stat Soc 27:191–203
Taillie C (1981) Lorenz ordering within the generalized gamma family of income distributions. In: Taillie C, Patil GP, Baldessari BA (eds) Statistical distributions in scientific work, vol 6. D Reidel Publishing Company, Dordrecht pp 181–192
Theil H (1967) Economics and information theory. North-Holland, Amsterdam
Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57:307–333
Wilfling B (1996a) Lorenz ordering of generalized beta-II income distributions. J Econ 71:381–388
Wilfling B (1996b) A sufficient condition for Lorenz ordering. Sankhya Ser B 58:62–69
Wilfling B, Krämer W (1993) The Lorenz ordering of Singh-Maddala income distributions. Econ Lett 43:53–57