Project no.
035086
Project acronym
EURACE
Project title
An Agent-Based software platform for European economic policy design with heterogeneous interacting agents: new insights from a bottom up approach to economic modelling and simulation

Instrument STREP

Thematic Priority IST FET PROACTIVE INITIATIVE “SIMULATING EMERGENT PROPERTIES IN COMPLEX SYSTEMS”

Deliverable reference number and title
D4.1: Empirical analysis of agents’ features distribution in real economies
Due date of deliverable: 29.02.2008

Actual submission date: __________

Start date of project: September 1st 2006 Duration: 36 months

Organisation name of lead contractor for this deliverable
Università Politecnica delle Marche - UPM

Revision 2

Project co-funded by the European Commission within the Sixth Framework Programme (2002–2006)		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	×
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

List of Figures

List of Tables

Abstract

While still a work in progress, this document focuses on the empirical analysis of real personal income data from a number of European countries for the EURACE project. Both descriptive and influential statistical issues are addressed. The document is going to be completed with the analysis of size and growth of European business firms.

Acknowledgements

This work was supported by the European Union through its 6th Framework Programme, FET-IST "Complex Systems". Funding for the STREP research project "EURACE" under contract no. 035086, is gratefully acknowledged. The author and contact person for this document is: Fabio Clementi - Department of Economics, Polytechnic University of Marche - Piazzale R. Martelli, 8, 60121, Ancona, Italy - Telephone: +39–071–22–07–103; fax: +39–071–22–07–102; e-mail address: `fabio.clementi@univpm.it`. The usual disclaimer applies.

1 The Shape of the Income Distribution[1]

A logical starting point for a discussion of the size distribution of incomes is Pareto's (1964, 1965) observation that the proportion of incomes above a level x is well approximated by

$$\bar{F}(x) = P_{>}(x) = \Pr(X > x) = \left(\frac{k}{x}\right)^{\alpha}, \quad k \leq x < \infty, \quad k, \alpha > 0, \tag{1}$$

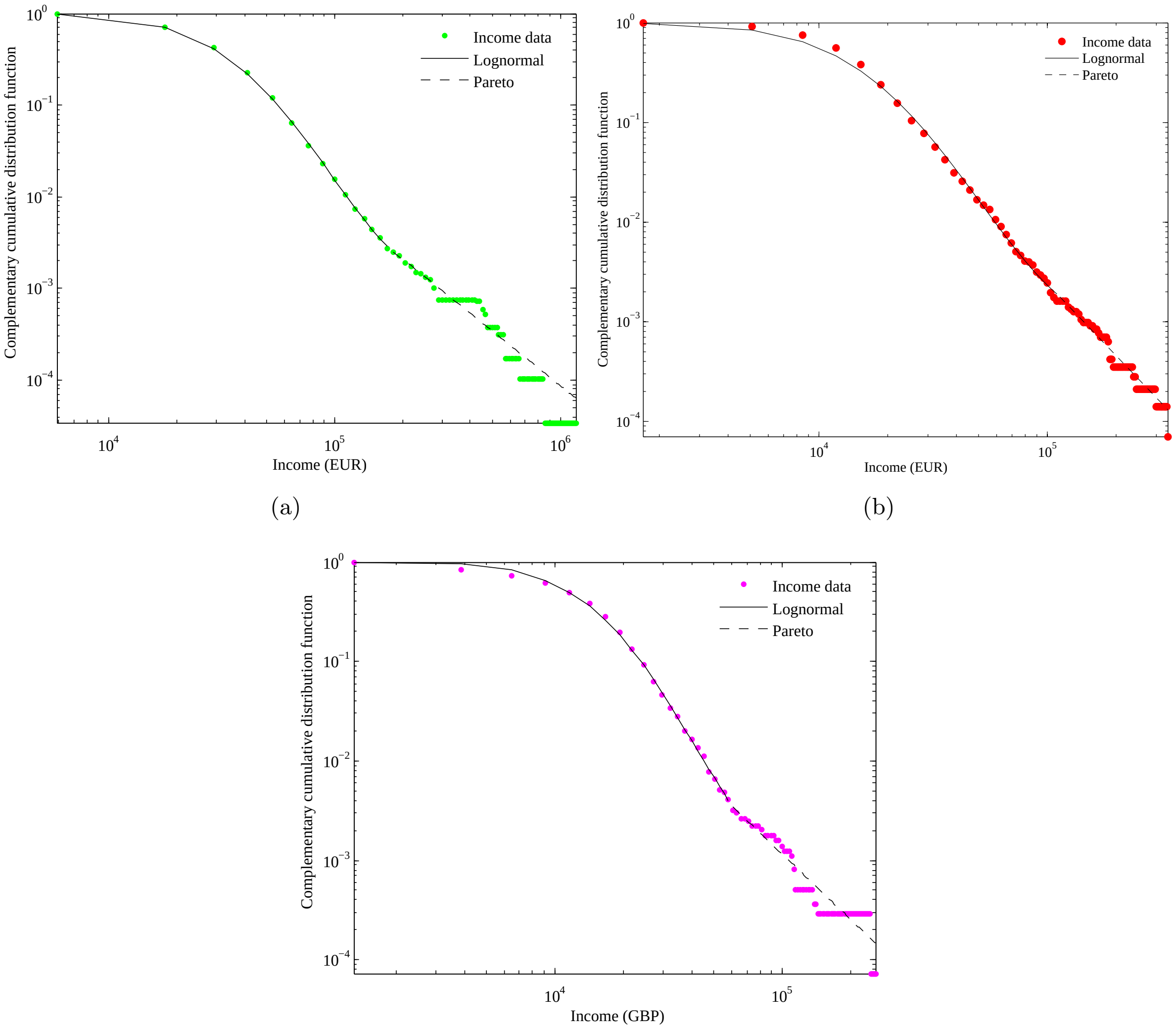
k being the minimum possible value of X. According to Pareto, the parameter α in (1), which turns out to be some kind of index of inequality of distribution, was usually not much different from 1.5. He asserted that there was some kind of underlying "law" that determined the form of income distributions. On occasion he even claimed that the value of α appeared to be invariant under changes of definition of income, changes due to taxation, etc., and to be insensitive to the choice of measuring individual or family income, or income per unit household member.

The classical Pareto's distribution (1) with its heavy tail soon became an accepted model for income. That is not to say that competitors did not abound. Gibrat's (1931) celebrated "law of proportional effect" leading to the lognormal distribution

$$P_{>}(x) = 1 - \Phi\left(\frac{\log x - \mu}{\sigma}\right), \quad 0 \leq x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0 \tag{2}$$

remains a strong competitor in any effort to fit income curves. The obvious advantage of the lognormal distribution is that following a simple transformation the enormous armature of inference for normal distributions is readily available. However, the functional form that is appropriate for modeling distributions depends on the definition of income and the particular part of the distribution in which one happens to be interested. For example, Aitchison & Brown (1954, 1957) argues that the lognormal hypothesis is particularly appropriate for the distribution of earnings in fairly homogeneous sections of the workforce, but when one examines the distribution of income from all sources he is quite likely to find in many instances that lognormality remains a reasonable assumption for the bulk of the income while the upper tail approximates more closely to the Pareto's distribution, as the evidence of Figure 1 bears out[2]. Moreover, empirically the values of parameter α are not "stable" indicating, for that part of the distribution where the Paretian approximation is suitable, that the distribution of income varies markedly over time (see Table 1 and Figure 2). The fact that the close fit of the Pareto's formula

[1] The discussion here focuses chiefly on income distribution, but it could be equally well extended to wealth



(a)

(b)

(c)

Figure 1: *Binned cumulative distribution of personal income in log-log scale for (a) Germany (2002), (b) Italy (2000) and (c) the United Kingdom (1998). The distribution is cumulated from the top. The fit to equation* (1) *is shown by the solid line, while the dashed line is the fit to equation* (2)*. The interpolations were performed via the Nonlinear Least-SQuares (NLSQ) method. The corresponding estimated parameter values are provided in Table 1.*

remains remarkably satisfactory only for high incomes rapidly led accumulating experience to point out that Pareto's law can hardly be expected to apply to the income distribution as a whole[3].

distribution, albeit with some care because of the distinctive features of wealth data. For a review of methods used to summarize and comparing wealth distributions, see *e.g.* Jenkins & Jäntti (2005).

[2] For the purposes of Figure 1, household income data with negative and zero values were eliminated and the remaining ones grouped into 100 equally spaced bins. The cumulative count of elements inside each bin was then plotted (on the horizontal axis) against the position of bin centers (on the vertical axis). In all cases, household income is the annual total post-tax-and-transfer household income, equivalized by dividing by the square root of the number of members in the household; it is expressed in nominal local currency units and were converted into 1995 constant prices by using the Consumer Price Index (CPI) obtained from the OECD (2003).

[3] The "Paretian upper tail" works well for the distribution of wealth. There is a superficial reason to suppose that a curve like Pareto's might be useful in this application. Wealth data are usually compiled with any accuracy

Table 1: *Estimated lognormal and Pareto's distribution parameters for the countries and years shown in Figures 1–2. Standard errors are listed in parentheses. Also shown are the estimates of Gibrat's index of inequality*[1], *measured as* $\beta = \frac{1}{\sigma\sqrt{2}}$.

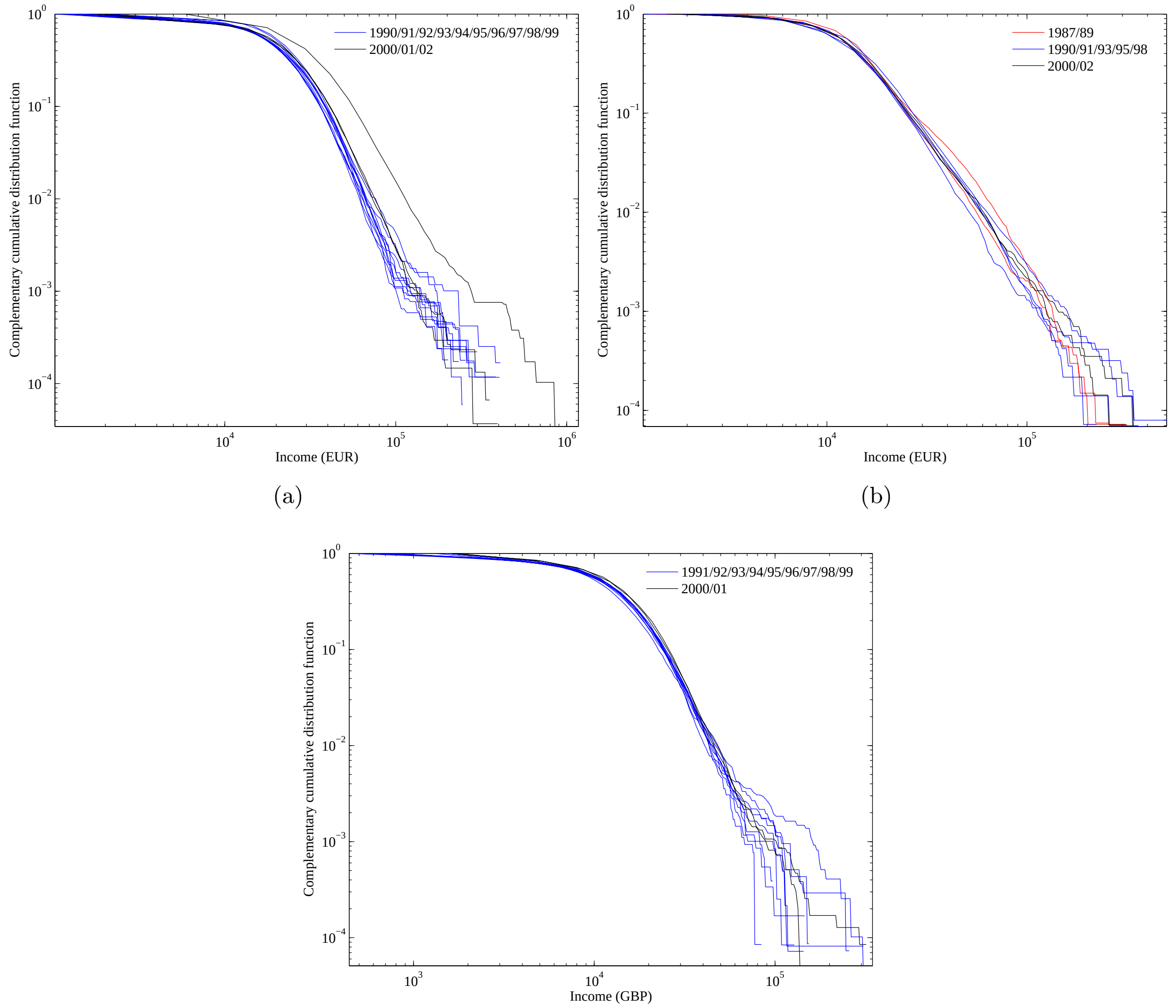
Country	Year	$\hat{\mu}$		$\hat{\sigma}$		$\hat{k}$		$\hat{\alpha}$		$\hat{\beta}$
GER	1990	5.07	(0.02)	0.52	(0.01)	9,464.02	(718.15)	2.35	(0.07)	1.37
	1991	5.19	(0.05)	0.58	(0.02)	11,632.73	(417.61)	2.67	(0.05)	1.22
	1992	5.26	(0.04)	0.63	(0.03)	1,121.38	(176.19)	1.55	(0.05)	1.12
	1993	5.35	(0.03)	0.69	(0.01)	6,541.44	(723.84)	2.39	(0.09)	1.03
	1994	5.28	(0.02)	0.65	(0.01)	5,763.27	(595.15)	2.23	(0.08)	1.08
	1995	5.35	(0.02)	0.69	(0.01)	3,609.71	(528.30)	1.93	(0.08)	1.02
	1996	5.60	(0.11)	0.79	(0.05)	2,298.85	(376.29)	1.76	(0.07)	0.89
	1997	5.38	(0.05)	0.70	(0.03)	18,975.12	(1,540.98)	3.79	(0.19)	1.00
	1998	5.55	(0.08)	0.77	(0.03)	15,754.00	(1,215.19)	3.37	(0.14)	0.92
	1999	5.53	(0.06)	0.76	(0.03)	6,872.99	(458.41)	2.29	(0.06)	0.93
	2000	5.46	(0.05)	0.74	(0.02)	14,394.27	(576.28)	3.03	(0.06)	0.95
	2001	5.39	(0.03)	0.70	(0.03)	16,378.20	(1,142.39)	3.41	(0.11)	1.01
	2002	5.78	(0.02)	0.72	(0.01)	17,094.20	(1,448.97)	1.97	(0.05)	0.98
ITA	1987	4.79	(0.03)	0.56	(0.02)	15,620.20	(171.27)	3.10	(0.03)	1.27
	1989	4.81	(0.02)	0.51	(0.02)	12,285.46	(193.75)	3.05	(0.03)	1.38
	1991	4.90	(0.02)	0.57	(0.01)	12,580.46	(252.66)	3.29	(0.05)	1.25
	1993	5.00	(0.03)	0.66	(0.02)	17,447.09	(152.81)	3.69	(0.02)	1.07
	1995	5.21	(0.06)	0.78	(0.04)	5,295.59	(424.83)	2.18	(0.06)	0.90
	1998	5.32	(0.04)	0.83	(0.03)	8,677.66	(578.48)	2.40	(0.05)	0.85
	2000	5.28	(0.04)	0.82	(0.03)	8,088.60	(477.31)	2.40	(0.05)	0.87
	2002	4.88	(0.01)	0.57	(0.01)	12,073.99	(125.45)	2.92	(0.02)	1.24
UK	1991	5.74	(0.11)	0.95	(0.04)	12,527.99	(654.26)	3.78	(0.14)	0.74
	1992	5.19	(0.07)	0.71	(0.04)	9,498.47	(520.88)	3.09	(0.09)	1.00
	1993	8.12	(0.98)	1.52	(0.20)	14,838.40	(315.70)	4.39	(0.10)	0.47
	1994	5.51	(0.10)	0.86	(0.04)	14,541.53	(252.71)	4.08	(0.06)	0.82
	1995	5.33	(0.07)	0.80	(0.03)	3,077.33	(395.67)	1.95	(0.08)	0.89
	1996	5.67	(0.11)	0.90	(0.04)	13,499.37	(479.98)	3.92	(0.10)	0.78
	1997	5.21	(0.04)	0.74	(0.02)	7,466.55	(758.72)	2.62	(0.12)	0.96
	1998	5.10	(0.04)	0.68	(0.03)	4,515.06	(373.22)	2.18	(0.06)	1.04
	1999	4.99	(0.03)	0.62	(0.02)	2,688.51	(258.90)	1.72	(0.05)	1.14
	2000	5.06	(0.03)	0.63	(0.02)	8,507.13	(659.65)	2.97	(0.10)	1.12
	2001	5.13	(0.05)	0.67	(0.03)	5,272.06	(267.53)	2.38	(0.04)	1.06

[1] While the Pareto's index provides a measure of the income inequality for the tail, the Gibrat's index provides a measure of the income inequality corresponding to the body of the distribution, and like the former is an inverse index of concentration: *i.e.*, if β has low values (large variance of the global distribution), the personal income is unevenly distributed; clearly, the reverse is true if β has high values. However it is worthwhile considering that that the measured values of Gibrat's index and Pareto's exponent are not consistent with the most widely used measures of income inequality (*e.g.* the Gini's coefficient) if one associates lower values of these indexes with increased inequality. In fact, if income follows either a lognormal or a Pareto's distribution throughout, then a clear correspondence can be found between the two measures and the preferred inequality index; however, observed distributions are not only lognormal and show a power-law only over a very limited range; hence, the correspondence breaks down (see *e.g.* Persky, 1992).

Source: author's own calculations based on the GSOEP-CNEF income data for the years 1990–2002, the SHIW income data for the years 1987–2002 and the BHPS-CNEF income data for the years 1991–2001.

This preliminary glimpse of evidence is perhaps sufficient to reinforce at least two conclusions

only for the moderately wealthy and above. Hence—excluding those whose wealth is unrecorded—one typically finds a single-tailed distribution. The Paretian property of the tail of the wealth distribution is demonstrated admirably by the Swedish data examined by Steindl (1965) where α is about 1.5 to 1.7.



(a)

(b)

(c)

Figure 2: *Time development of the personal income distribution for (a) Germany (1990–2002), (b) Italy (1987–2002) and (c) the United Kingdom (1991–2001).*

that may have suggested themselves earlier in the discussion.

- Neither the Pareto's nor the lognormal hypothesis provides a "law" of distribution in the strict sense that it (or a particular member of either family) is an exact model of income distribution. In particular, it is nonsense to suppose that the Pareto's curve (where applicable) should remain stable over time: as it happens, α fluctuates within narrow bounds.

- Nevertheless, one or other functional form is a reasonable approximation of the shape of income distribution when one looks at a well-defined piece of it.

This could be the end of the story, except for two nagging questions. First, since the Pareto's law seems to hold only for the upper tail of the income distribution, how might one determine the cutoff point above which it could be expected to hold? And second, what kind of model would account for income distribution throughout its entire range? Subject-specific results are discussed in what follows.

2 On Bootstrapping to "Endogenize" Tail Estimation

Inference procedures for the classical Pareto's (power-law) distribution have been discussed extensively in the literature (see *e.g.* Arnold, 1983, Johnson et al., 1994, Kleiber & Kotz, 2003, and Quandt, 1996, for a considerable in-depth discussion). Perhaps the oldest and still among the most popular technique for estimating parameters relies on the observation that the logarithm of the survival function (1) is linear, *i.e.*

$$\log \bar{F}(x) = \alpha \log k - \alpha \log x.$$

Fitting a straight line by least squares leads to the following regression estimator of α

$$\hat{\alpha} = \frac{n \sum_{i=1}^{n} \log x_i \sum_{i=1}^{n} \log \bar{F}(x_i) - \sum_{i=1}^{n} \log x_i \log \bar{F}(x_i)}{n \sum_{i=1}^{n} (\log x_i)^2 - \left(\sum_{i=1}^{n} \log x_i \right)^2},$$

while an estimator for k can be obtained by exploiting the mathematical relationship

$$\hat{k} = \text{antilog}\left(\frac{\hat{C}_{\hat{\alpha}}}{\hat{\alpha}} \right),$$

where $\hat{C}_{\hat{\alpha}} = \overline{\log \bar{F}(x)} + \hat{\alpha}\overline{\log x}$ is the regression constant estimate.

Unfortunately, this approach is not immune from objections (Aigner & Goldberger, 1970; Clauset et al., 2007; Coronel-Brizio & Hernández-Montoya, 2005; Goldstein et al., 2004; Sornette, 2004; Weron, 2001), and some alternative methods for estimating the parameters of a power-law distribution that are generally more accurate and robust have been proposed. Among these, the maximum likelihood estimator of α introduced by Hill (1975)—which is known to be asymptotically normal (Hall, 1982) and consistent (Mason, 1982)—does not assume a parametric form for the entire distribution function, but focuses only on the tail behavior. That is, if $x_{[n]} \geq x_{[n-1]} \geq \ldots \geq x_{[n-m]} \geq \ldots \geq x_{[1]}$, with $x_{[i]}$ denoting the i^{th} order statistic, are the sample elements put in descending order, then the Hill's estimator for α based on the m largest order statistics is

$$\hat{\alpha}_n(m) = \left[\frac{1}{m} \sum_{i=1}^{m} \left(\log x_{n-i+1} - \log x_{n-m} \right) \right]^{-1}, \tag{3}$$

where n is the sample size and m an integer value in $[1, n]$. Unfortunately, it is difficult to choose the right value of m. In practice, $\hat{\alpha}_n(m)$ is plotted against m and one looks for a region where the plot levels off to identify the optimal sample fraction to be used in the estimation of α (Embrechts et al., 1997; Resnick, 1997). Moreover, the finite-sample properties of the estimator (3) depend crucially on the choice of m: increasing m reduces the variance because more data are used, but it increases the bias because the power-law is assumed to hold only in the extreme tail.

Over the last twenty years, estimation of the Pareto's index α has received considerable attention in extreme value statistics (see *e.g.* Lux, 2001). All of the proposed estimators, including the Hill's estimator, are based on the assumption that the number of observations in the upper tail to be included, m, is known. In practice, m is unknown; therefore, the first task is to identify which values are really extreme values. Tools from exploratory data analysis, as the quantile-quantile plot and/or the mean excess plot, might prove helpful in detecting graphically the quantile $x_{[n-m]}$ above which the Pareto's relationship is valid; however, they do not propose any formal computable method and, imposing an arbitrary threshold, they only give very rough estimates of the range of extreme values.

Given the bias-variance trade-off for the Hill's estimator, a general and formal approach in determining the best m value is the minimization of the *Mean Squared Error* (MSE) between $\hat{\alpha}_n(m)$ and the theoretical value α. Unfortunately, in empirical studies of data the theoretical value of α is not known. Therefore, an attempt to find an approximation to the sampling distribution of the Hill's estimator is required. To this end, a number of innovative techniques in the statistical analysis of extreme values proposes to adopt the powerful bootstrap tool to find the optimal number of order statistics adaptively (Dacorogna et al., 1992; Danielsson et al., 2001; Hall, 1990; Lux, 2000). By capitalizing on these recent advances in the extreme value statistics literature, in this section a subsample semi-parametric bootstrap algorithm is proposed in order to make a more automated, "data-driven" selection of the extreme quantiles useful for studying the upper tail of income distribution, and to end up at less ambiguous estimates of α[4]. This methodology is described in Section 2.1, and its application to Italian income data is given in Section 2.2.

2.1 Estimation Technique for Threshold Selection

To find the optimal threshold k_n^*—or equivalently the optimal number m^* of extreme sample values above that threshold—to be used for estimation of α, the MSE of the Hill's estimator (3) is minimized for a series of thresholds $k_n = x_{[n-m]}$, and the k_n value at which it attains the minimum is picked as k_n^*. Given that different threshold series choices define different sets of possible observations to be included in the upper tail of a specific observed sample $\mathbf{x}_n = \{x_i; i = 1, 2, \ldots, n\}$, only the observations exceeding a certain threshold that additionally follow a $\mathrm{Par}\left(k_n, \hat{\alpha}_n(m)\right)$ distribution—where $\hat{\alpha}_n(m)$ is a prior estimate for each threshold k_n of the Pareto's tail index obtained through the Hill's statistic—are included in the series.

In order to check this condition, a (two-sided) *Kolmogorov-Smirnov* (*K-S*) goodness-of-fit test is performed for each threshold in the original sample for the null hypothesis

$$H_0 : \hat{F}_n(y) = F(y)$$

versus the general alternative of the form

$$H_1 : \hat{F}_n(y) \neq F(y),$$

where Y is a standard exponential variable, *i.e.* $p_y(y) = e^{-y}$, $y > 0$, and $\hat{F}_n(y)$ is the empirical cumulative distribution function[5]. The formal steps in making a test of H_0 are as follows (Stephens, 1970, 1974; D'Agostino & Stephens, 1986):

[4] Hill himself devised a data-analytic method for choosing m which is based on sequentially testing appropriate functions of the observations for exponentiality. However, as observed by Hall & Welsh (1985), the exponential approximation deteriorates very gradually, leading Hill's method to largely overestimate m, and thus α by Equation (3).

[5] A basic distributional property of the standard Pareto's distribution is its close relationship with the exponential distribution. Indeed, using the rule of transformation of random variables

$$p_y(y) = p_x\left[f^{-1}(y)\right]\left|\frac{\mathrm{d}\, f^{-1}(y)}{\mathrm{d}\, y}\right|,$$

where $y = f(x) = \alpha \log\left(\frac{x}{k}\right)$ and $x = f^{-1}(y) = ke^{\frac{y}{\alpha}}$, one immediately gets that Y has a standard exponential distribution, *i.e.* an exponential distribution with parameter equal to 1. Therefore, hypothesis testing for the classical Pareto's distribution may follow immediately from the exponential case.

1. Calculate the original K-S test statistic, D, by using the formula

$$D = \sup_{-\infty<y<\infty} \left| \hat{F}_n(x) - F(y) \right|.$$

2. Calculate the modified form, T^*, by using the formula

$$T^* = D\left(\sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}\right). \qquad (4)$$

3. Reject H_0 if T^* exceeds the cutoff level, z, for the chosen significance level.

To obtain an estimate of finite-sample bias and variance (and thus MSE) at each threshold coming from the null hypothesis H_0, a natural criterion is to use the *bootstrap* (Davison & Hinkley, 1997; Efron & Tibshirani, 1993). In its purest form, the bootstrap involves approximating an unknown distribution function, $F(x)$, by the empirical distribution function, $\hat{F}_n(x)$. However, most times the empirical distribution model from which one resamples in a purely nonparametric bootstrap is not a good approximation of the distribution shape in the tail[6]. Therefore, the tail data are initially smoothed by fitting a Pareto's cumulative distribution function

$$PD_{\hat{\alpha}_{n_1},\hat{k}_{n_1}}(x) = p = 1 - \left(\frac{k_{n_1}}{x}\right)^{\hat{\alpha}_{n_1}} \qquad (5)$$

to the $n_1 \leq n$ observations $\mathbf{x}_{n_1} = \{x \in \mathbf{x}_n : T^* \leq z\}$, with $\hat{k}_{n_1} = \min_i x_i$ and $\hat{\alpha}_{n_1} = \frac{n_1}{\sum_{i=1}^{n_1}\left(\log x_i - \log \hat{k}_{n_1}\right)}$ being the maximum likelihood estimators for the Pareto's distribution parameters, and then the bootstrap samples are drawn from the set of quantiles $\mathbf{x}_{n_1}^p = \left\{x \in \mathbf{x}_{n_1} : PD_{\hat{\alpha}_{n_1},\hat{k}_{n_1}}(x) \geq p\right\}$ obtained directly from inverting the estimated model (5).

The adopted methodology can be summarized as follows:

1. Evaluate the estimate $\hat{\alpha}_n(m)$ of the Pareto's tail index for each threshold in the original sample $\mathbf{x}_n$ by using the Hill's estimator (3).

2. For each threshold in the original sample, make the log-transformation using the parameters estimated from the data and test the Pareto's approximation by computing the value of the *K-S* test statistic (4).

3. Fit the model (5) to the subset of data $\mathbf{x}_{n_1}$ belonging to the null hypothesis H_0.

4. Select B independent bootstrap samples $\mathbf{x}_1^{\#}, \mathbf{x}_2^{\#}, \ldots, \mathbf{x}_B^{\#}$, each consisting of n_1 values drawn with replacement from the set of quantiles $\mathbf{x}_{n_1}^p$ obtained by inverting the fitted model (5).

5. For each bootstrap sample $\mathbf{x}_b^{\#}$, $b = 1, 2, \ldots, B$, and for each threshold $k_{n_1}^{\#}$ in the bootstrap sample, evaluate the bootstrap estimate $\hat{\alpha}_{n_1}^{\#}(m_1)$ of the Pareto's tail index by using the Hill's estimator (3).

[6] When income micro-data come from a sample survey, there is usually a problem of *non-response*, especially at the upper extremes of the income distribution. For example, recipients of large amounts of investment income, representing a substantial proportion of the total for investment income, tend to be a small group of individuals concentrated in the upper end of the distribution more likely not to respond when sampled for a survey. This non-response bias can thus give rise to a significant departure from the true underlying model for the largest observations. See The Canberra Group (2001) on this and other conceptual and methodological problems for data on income distribution.

6. For each threshold $k^{\#}_{n_1}$, calculate the bias

$$\left\{E\left[\hat{\alpha}^{\#}_{n_1}(m_1)\right] - \hat{\alpha}_n(m)\right\}_B = \frac{1}{B}\sum_{b=1}^{B}\hat{\alpha}^{\#}_{n_1,b}(m_1) - \hat{\alpha}_n(m),$$

the variance

$$\hat{se}^2_B = \frac{1}{B-1}\sum_{b=1}^{B}\left[\hat{\alpha}^{\#}_{n_1,b}(m_1) - \frac{1}{B}\sum_{c=1}^{B}\hat{\alpha}^{\#}_{n_1,c}(m_1)\right]^2, \tag{6}$$

and the mean squared error

$$M\hat{S}E_B = \left\{E\left[\hat{\alpha}^{\#}_{n_1}(m_1)\right] - \hat{\alpha}_n(m)\right\}^2_B + \hat{se}^2_B$$

of the Hill's tail index estimates.

7. Select as the optimal threshold $k^*_n = x_{[n-m^*]}$ that threshold where the MSE attains its minimum.

Minimizing the MSE, thus, amounts to find the MSE minimizing number of order statistics $m^* = \arg\min\limits_{m} MSE$ from which one infers the optimal estimate of the tail index $\hat{\alpha}^*_n(m^*)$.

2.2 Empirical Application: The Italian Personal Income Distribution

The data used to illustrate how the methodology proposed in Section 2.1 can be applied to the analysis of income distribution have been selected from the "Survey on Household Income and Wealth" (SHIW) provided by the Bank of Italy.

Figure 3(a) depicts the outcomes of the complete sequences of K-S test for a selection of tail fractions. Blue points mark all the observations for which the modified K-S statistic (4) does not exceed the 5% cutoff level $z = 1.358$ (see solid lines)[7]. The figure indicates the tail regions that may be tentatively regarded as appropriate for the implementation of the semiparametric bootstrap technique.

The Hill's estimator (3) is reported in Figure 3(b) for tail $\leq 25\%$ of the full sample size. In the figure, the optimal number of extreme sample values is reported, namely $m^* = 3,222$, providing the following value for the tail power-law exponent: $\hat{\alpha}^*_n(m^*) = 2.50 \pm 0.08$, where the error (with 95% confidence) has been obtained through the jackknife method (Dacorogna et al., 1992; Pictet et al., 1996)[8]. In the computation, 974 resamples have been used, and the subsample size has been set equal to the number of observations not rejected by the K-S test at the 5% level (see Section 2.1 and Figure 3(a))[9].

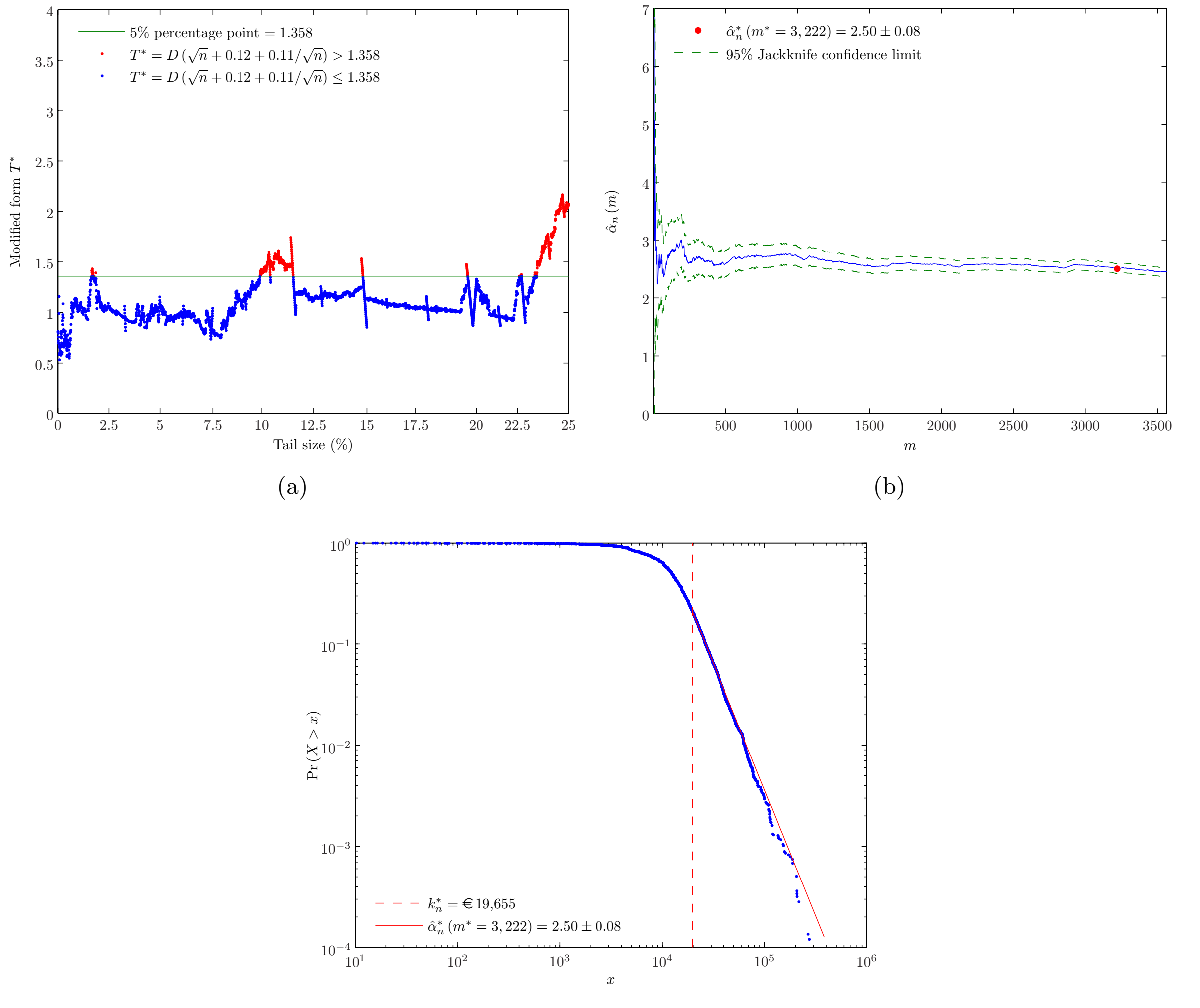
[7] The 5% significance point $z = 1.358$ with which the test has been conducted comes from Table 1 in Stephens (1970); see also Table 1A in Stephens (1974).

[8] The procedure known as a *jackknife* forms a matrix of size $[(g-1)\times h]\times g$ from the vector of data points of length $g \times h$. The vector is first divided into g blocks of size h. Each column of the matrix is formed by deleting a block from the vector. Here the standard version of the jackknife ($h = 1$) is adopted. On average, the jackknife error is slightly higher than the theorerical variance estimate. However, the error obtained through the jackknife method better relfects the data set that is actually used. The theoretical error does not account for the degree of noise produced by the resampling, while the jackknife method does.

[9] The number of bootstrap replications, B, has been determined according to the method recently developed by Andrews & Buchinsky (2000). Their approach is based on the following asymptotic result

$$\sqrt{B}\frac{\hat{se}_B - \hat{se}_\infty}{\hat{se}_\infty} \xrightarrow{d} N\left(0, \frac{2+\gamma}{4}\right), \tag{7}$$

where $\hat{se}_\infty$ is the "ideal" bootstrap standard error estimator of $\hat{\alpha}_n(m)$ based on an infinite number of bootstrap



(a)

(b)

(c)

Figure 3: *The Italian personal income distribution in 2000. (a) Modified K-S statistic* (4) *as a function of the tail size. (b) The Hill's estimator* (3)*. The dashed lines represent the 95% confidence limits of the tail index estimates computed by using the jackknife method. The point marks the optimal number of extreme sample values m^*. (c) Complementary cumulative distribution and power-law fit by using the estimated optimal value for α.*

The use of this $\hat{\alpha}_n^*$ optimal value produces the fit shown by the solid lines in Figure 3(c),

repetitions—if they were costless in terms of computational costs associated with them, $\hat{se}_B$ is the bootstrap standard error estimator given by the square root of (6) based on B bootstrap repetitions, and γ is a measure of the excess kurtosis of the bootstrap distribution of the Hill's estimator. Andrews and Buchinsky's method requires to specify two parameters. The *Percentage Deviation Bound* (PDB) indicates the maximum desired percentage difference between $\hat{se}_B$ and $\hat{se}_\infty$. The second parameter, τ, defines the *probability* with which $\hat{se}_B$ and $\hat{se}_\infty$ differ by more than PDB percent; that is

$$\Pr\left(100\frac{|\hat{se}_B - \hat{se}_\infty|}{\hat{se}_\infty} \le PDB\right) = 1 - \tau.$$

Therefore, given the result shown in (7), one can write

$$\Pr\left(100\frac{|\hat{se}_B - \hat{se}_\infty|}{\hat{se}_\infty} \le 50 z_{1-\frac{\tau}{2}}\sqrt{\frac{2+\gamma}{B}}\right) \xrightarrow{p} 1 - \tau,$$

where the complementary cumulative distribution is plotted on a double logarithmic scale. The vertical dashed line indicates the optimal value of the threshold parameter attained by subsample semi-parametric bootstrapping, that is $k_n^* =$ €19,655. As one can see, the procedure succeeds in avoiding deviations from linearity for the largest observations that might strongly influence the estimation of α, illustrating therefore the importance of optimally choosing the tail threshold.

3 An Overall Description of Income Distribution

As discussed in Section 1, it has long been known that the Pareto's distribution, providing a description of the density for income values above some lower bound $k > 0$, is usually unsuitable to approximate the full distribution of income. Indeed, even though there are simple expressions for the moments which depend only on the Pareto's parameters α and k and the expressions for most common inequality measures depend only on α—so that the (inverse of) α may also be considered as an inequality measure, the apparent attractions of the Pareto's distribution evaporate somewhat when one considers its implications for the distribution of income amongst the population as a whole, *i.e.* including units with income less than k. For example, Atkinson & Harrison (1978) show how expressions for the Gini's coefficient and the relative mean deviation depend on assumptions about the size of "excluded population" (*i.e.* the proportion of the population with income below k) and its average income. In particular, α no longer has such a straightforward interpretation. For example, an increase in α may be associated with an increase in inequality according to the Gini, but a decrease according to the Coefficient of Variation.

This suggests fitting of parametric models for the distribution of income as a whole. The income distribution literature suggests a large number of candidates (see *e.g.* the comprehensive survey by Kleiber & Kotz, 2003). 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's distribution in Champernowne, 1953). Another approach derives

or

$$PDB \doteq 50 z_{1-\frac{\tau}{2}} \sqrt{\frac{2+\gamma}{B}},$$

so that

$$B \approx 2,500 z^2_{1-\frac{\tau}{2}} \frac{2+\gamma}{PDB^2}, \tag{8}$$

where $z^2_{1-\frac{\tau}{2}}$ is the $1-\frac{\tau}{2}$ quantile of the standard normal distribution. Because γ is unknown, Andrews and Buchinsky propose a three-step method for choosing B:

1. Set $\gamma = 0$ and evaluate (8) to obtain an initial estimate B_1 of B. Setting $\gamma = 0$ corresponds to assuming that the bootstrap distribution of the Hill's estimator has no excess kurtosis.
2. Perform B_1 bootstrap replications, and then compute the bootstrap bias-corrected estimator of γ

$$\hat{\gamma}_B = \frac{\frac{1}{B-1}\sum_{b=1}^{B}\left[\hat{\alpha}^{\#}_{n1,b}(m_1) - \frac{1}{B}\sum_{c=1}^{B}\hat{\alpha}^{\#}_{n1,c}(m_1)\right]^4}{\left\{\frac{1}{B-1}\sum_{b=1}^{B}\left[\hat{\alpha}^{\#}_{n1,b}(m_1) - \frac{1}{B}\sum_{c=1}^{B}\hat{\alpha}^{\#}_{n1,c}(m_1)\right]^2\right\}^2} - 3.$$

3. Obtain a second estimate B_2 of B by using $\gamma = \hat{\gamma}_B$ in (8). If $B_1 \geq B_2$, take $B^* = \max(B_1, B_2) = B_1$ as the desired number of bootstrap repetitions; otherwise, draw a further $B^* - B_1$ bootstrap samples.

In the bootstrapping application, the following pair (PDB, τ) of parameters is used: $PDB = 5$ and $\tau = 0.05$. Then, $z_{1-\frac{\tau}{2}} = 1.96$, and from step 1 of the Andrews and Buchinsky's method the estimated initial number of replications needed is 768. With this set of bootstrap replications in hand, the revised number based on steps 2 and 3 is 974, with a further 206 bootstrap samples needed.

flexible analytical forms by considering solely their ability to ensure a satisfactory fit to empirical data (*e.g.* the gamma density of Salem & Mount, 1974, and the generalized beta models of McDonald, 1984). Finally, models are also derived from differential equations specified to capture regularity features of observed income distributions (*e.g.* the models proposed by Singh & Maddala, 1976, and Dagum, 1977).

The families of two-parameter models are evidently limited in the variety of shapes of income distributions that they can be expected to describe[10]. One way forward is to consider extensions to the basic forms to make them more flexible[11]. Several other families of distributions have been shown to have merit in capturing some important features of the distribution—*e.g.* the cyclical movements in the observed income distribution, and thus the impact of macroeconomic factors on the distribution of income (Metcalf, 1969); many of these functional forms are interrelated, in the sense that one is a special form of another, or one approximates another asymptotically. In the light of these considerations, the present section proposes a three-parameter distribution that is a generalization of the Pareto's and the Weibull distribution using a new approach recently advanced by Kaniadakis (2001, 2002, 2005) to describe physical systems, and already exploited for statistical modeling in econometrics (Rajaonarison et al., 2005)[12]. This approach characterizes the distribution as solution of a maximum entropy model based on the κ-deformed exponential and logarithmic functions

$$\exp_\kappa(x) = \left(\sqrt{1+\kappa^2x^2}+\kappa x\right)^{1/\kappa}, \quad x \in \mathbf{R}, \tag{9a}$$

$$\log_\kappa(x) = \frac{x^\kappa - x^{-\kappa}}{2\kappa}, \quad x \in \mathbf{R}^+. \tag{9b}$$

These κ-deformed functions satisfy most properties of the standard exponential and logarithm, which are recovered as the real deformation parameter κ approaches zero; for applications to statistical analysis of income distribution, the most interesting property is their power-law asymptotic behavior

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

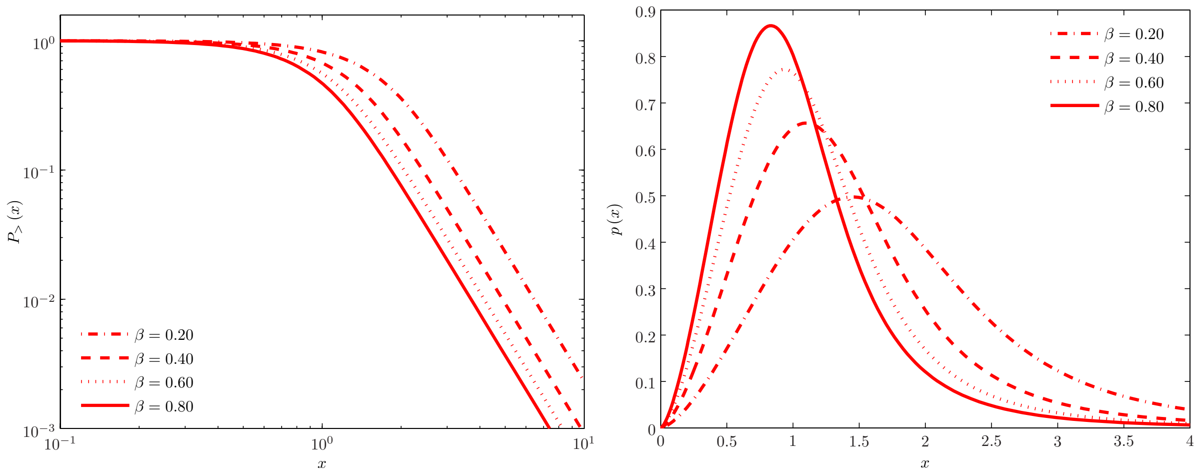
and thus their ability to satisfy the weak Pareto's law (Mandelbrot, 1960)[13].

[10] As already noted in Section 1, if one considers the entire range of income the performance of the lognormal model in the upper end is far from being satisfactory, whereas the fit provided by the Pareto's distribution appears distinctly superior. In terms of goodness-of-fit, the gamma distribution outperforms the lognormal at the two tails of the distribution (McDonald & Ransom, 1979), even though in the middle income range it overcorrects for the positive skewness of the data (Majumder & Chakravarty, 1990).

[11] For example, Montroll & Shlesinger (1982, 1983) shows that a mixture of lognormal distributions with a geometric weighting distribution would have essentially a lognormal main part but a Pareto-type distribution in the upper tail. A generalization of the model above introduced by Reed & Jorgensen (2004) assumes that a good fit of the whole range of incomes is provided by a distribution exhibiting Paretian (power-law) behavior in both tails and therefore referred to as the "double Pareto-lognormal" distribution. This distribution arises as that of the state of a geometric Brownian motion with lognormally distributed initial state after an exponentially distributed length of time.

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

[13] The use of the entropy concept in the analysis of income distribution is not new. For example, Ord et al. (1981) pointed out that the Pareto's, gamma and lognormal distributions might be selected if one uses a criterion of maximum entropy (different measures of entropy of course lead to different maximizing distributions). On the inequality side, Dalton's (1920) "Principle of Population" paved the way of introducing a general measure of inequality that led to the notion of entropy-like function much earlier to the works of Shannon (1948) in information theory. This suggestion then found expression in the entropy-based measure of inequality proposed by Theil (1967), which naturally contributed to the development of a general information-theoretic approach to the measurement of inequality (Cowell, 1980a,b, and Cowell & Kuga, 1981a,b).



(a) (b)

Figure 4: *(a) Plot of the κ-generalized CCDF given by Equation (10) versus x for some different values of $\beta\,(= 0.20, 0.40, 0.60, 0.80)$, and fixed $\alpha\,(= 2.50)$ and $\kappa\,(= 0.75)$. (b) Plot of the κ-generalized PDF given by Equation (11) versus x for some different values of $\beta\,(= 0.20, 0.40, 0.60, 0.80)$, and fixed $\alpha\,(= 2.50)$ and $\kappa\,(= 0.75)$. Notice that the distribution spreads out (concentrates) as the value of β decreases (increases).*

3.1 The κ-Generalized Distribution: Definitions and Interrelations

The κ-generalized Complementary Cumulative Distribution Function (CCDF) is given by

$$P_>(x) = \exp_\kappa(-\beta x^\alpha), \quad x \in \mathbf{R}^+, \tag{10}$$

being $P_>(x)$ the probability of finding the distribution variable X with a value greater than x. The income variable x is defined as $x = \frac{z}{\langle z \rangle}$, being z the absolute personal income and $\langle z \rangle$ its mean value. Then the dimensionless variable x represents the personal income in units of $\langle z \rangle$. The constant $\beta > 0$ is a characteristic scale, since its value determines the scale of the probability distribution: if β is large, then the distribution will be more concentrated; if β is small, then it will be more spread out (see Figure 4(a)–(b)). The exponent $\alpha > 0$ quantifies the curvature (shape) of the distribution, which is less (more) pronounced for lower (higher) values of the parameter, as seen in Figure 5(a)–(b). Finally, as one can observe in Figure 6(a)–(b), the deformation parameter $\kappa \in [0, 1)$ measures the fatness of the upper tail: the larger (smaller) its magnitude, the fatter (thinner) the tail.

The function $P_>(x)$ defined through Equation (10) can be viewed as a generalization of the ordinary stretched exponential (Laherrère & Sornette, 1998; Sornette, 2004), *i.e.* $P_>^0(x) = \exp(-\beta x^\alpha)$, which recovers in the $\kappa \to 0$ limit. It is remarkable that $P_>(x)$ for $x \to 0^+$ behaves as the ordinary stretched exponential

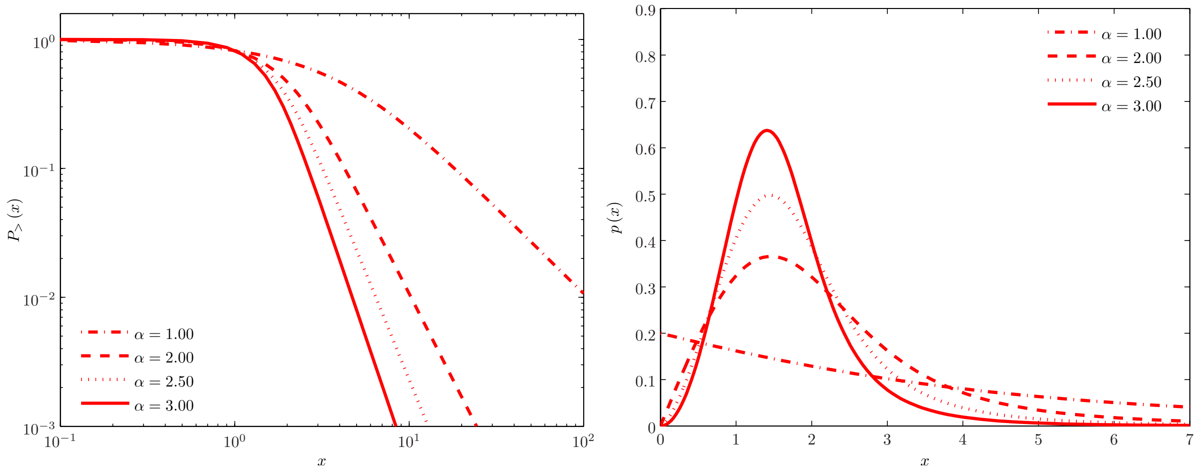
$$P_>(x) \underset{x \to 0^+}{\sim} \exp(-\beta x^\alpha),$$

while for $x \to \infty$ presents a power-law tail

$$P_>(x) \underset{x \to +\infty}{\sim} (2\beta\kappa)^{-\frac{1}{\kappa}} x^{-\frac{\alpha}{\kappa}}.$$

The Probability Density Function (PDF), $p(x) = -\frac{\mathrm{d}P_>(x)}{\mathrm{d}x}$, is given by

$$p(x) = \frac{\alpha\beta x^{\alpha-1} \exp_\kappa(-\beta x^\alpha)}{\sqrt{1 + \beta^2\kappa^2 x^{2\alpha}}}, \tag{11}$$



(a)

(b)

Figure 5: *(a) Plot of the κ-generalized CCDF given by Equation* (10) *versus x for some different values of $\alpha\,(= 1.00, 2.00, 2.50, 3.00)$, and fixed $\beta\,(= 0.20)$ and $\kappa\,(= 0.75)$. (b) Plot of the κ-generalized PDF given by Equation* (11) *versus x for some different values of $\alpha\,(= 1.00, 2.00, 2.50, 3.00)$, and fixed $\beta\,(= 0.20)$ and $\kappa\,(= 0.75)$. Notice that the curvature (shape) of the distribution becomes less (more) pronounced when the value of α decreases (increases). The case $\alpha = 1.00$ corresponds to the ordinary exponential function.*

and can be viewed as a generalization of the Weibull distribution (*e.g.* Johnson et al., 1994), *i.e.* $p^0(x) = \alpha\beta x^{\alpha-1}\exp(-\beta x^\alpha)$, which recovers in the $\kappa \to 0$ limit. The function $p(x)$ given by Equation (11) for $x \to 0^+$ behaves as a Weibull distribution

$$p(x) \underset{x\to 0^+}{\sim} \alpha\beta x^{\alpha-1}\exp(-\beta x^\alpha),$$

while for $x \to +\infty$ reduces to the Pareto's law

$$p(x) \underset{x\to +\infty}{\sim} \frac{\alpha}{\kappa}(2\beta\kappa)^{-\frac{1}{\kappa}} x^{-\left(\frac{\alpha}{\kappa}+1\right)}.$$

Using the complementary relation $P_\leq(x) = 1 - P_>(x)$, the quantile function is immediately obtained as

$$x = P_\leq^{-1}(u) = \beta^{-\frac{1}{\alpha}}\left[\log_\kappa\left(\frac{1}{1-u}\right)\right]^{\frac{1}{\alpha}}, \quad 0 < u < 1,$$

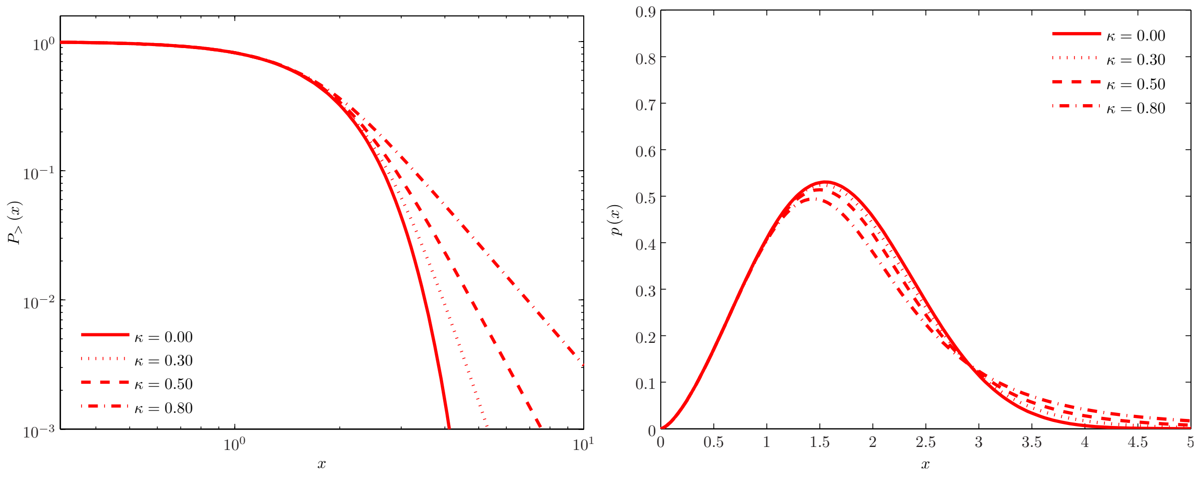
from which follows that the median of the distribution is

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

The mode is at

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

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



(a)

(b)

Figure 6: *(a) Plot of the κ-generalized CCDF given by Equation* (10) *versus x for some different values of $\kappa\,(= 0.00, 0.30, 0.50, 0.80)$, and fixed $\beta\,(= 0.20)$ and $\alpha\,(= 2.50)$. (b) Plot of the κ-generalized PDF given by Equation* (11) *versus x for some different values of $\kappa\,(= 0.00, 0.30, 0.50, 0.80)$, and fixed $\beta\,(= 0.20)$ and $\alpha\,(= 2.50)$. Notice that the upper tail of the distribution fattens (thins) as the value of κ increases (decreases). The case $\kappa = 0.00$ corresponds to the ordinary stretched exponential (Weibull) function (Johnson et al., 1994; Laherrère & Sornette, 1998; Sornette, 2004).*

3.2 Moments and Related Parameters

The r^{th}-order moment about the origin of the κ-generalized distribution equals

$$\mu_r^{'} = \int\limits_0^\infty x^r p(x)\,\mathrm{d}\,x = \frac{(2\beta\kappa)^{-\frac{r}{\alpha}}}{1+\frac{r}{\alpha}\kappa}\frac{\Gamma\left(\frac{1}{2\kappa}-\frac{r}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{r}{2\alpha}\right)}\Gamma\left(1+\frac{r}{\alpha}\right), \tag{12}$$

where $\Gamma(x)$ is the Euler's Gamma function $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\mathrm{d}t$. Specifically, $\mu_1^{'} = m$ is the mean of the distribution.

A formula for the variance is obtained by converting Equation (12) to the moment about the mean using the general equation $\mu_r = \sum_{j=0}^{r}\binom{r}{j}(-1)^{r-j}\mu_j^{'}m^{r-j}$; hence, for $r = 2$ one gets

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

In this way it is also possible to define the standardized moments of the distribution, which are in turn used to define skewness and excess kurtosis, respectively given by

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3^{'} - 3m\sigma^2 - m^3}{\sigma^3}$$

and

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4^{'} - 3\sigma^4 - 4\gamma_1\sigma^3 m - 6\sigma^2 m^2 - m^4}{\sigma^4}.$$

3.3 Lorenz Curve and Inequality Measures

The practical advantage of having the quantile function available in closed form consists in the derivation of Lorenz-ordering results. In statistical terms, for any general distribution supported on the nonnegative half-line with a finite and positive first moment the Lorenz (1905) curve can be written in the form $L(u) = \frac{\int_0^u P_{\leq}^{-1}(t)\,\mathrm{d}\,t}{\int_0^1 P_{\leq}^{-1}(u)\,\mathrm{d}\,u}$, $u \in [0,1]$, where $m = \int_0^1 P_{\leq}^{-1}(u)\,\mathrm{d}\,u$ is the quantile formula for the mean and $P_{\leq}^{-1}(u)$ is the quantile function (Gastwirth, 1971). Hence, the Lorenz curve for the κ-generalized distribution is defined by

$$\begin{aligned} L_\kappa(u) =& 1 - \frac{1+\frac{\kappa}{\alpha}}{2\Gamma\left(\frac{1}{\alpha}\right)} \frac{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right)} \left\{ 2\alpha\,(2\kappa)^{\frac{1}{\alpha}}(1-u)\left[\log_\kappa\left(\frac{1}{1-u}\right)\right]^{\frac{1}{\alpha}} + \right. \\ & \left. + B_X\left(\frac{1}{2\kappa}-\frac{1}{2\alpha},\frac{1}{\alpha}\right) + B_X\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}+1,\frac{1}{\alpha}\right) \right\}, \end{aligned} \tag{13}$$

where $B_x(s,r)$ is the incomplete Beta function given by $B_x(s,r) = \int_0^x t^{s-1}(1-t)^{r-1}\,\mathrm{d}t$ with $X = (1-u)^{2\kappa}$.

Also, according to the results of Section 3.2 several measures of inequality can be considered. First, the Gini's (1914) measure of income inequality can be derived using the representation in terms of order statistics $G = 1 - \frac{1}{m}\int_0^\infty \left[1 - P_{\leq}(x)\right]^2 \mathrm{d}\,x$ due to Arnold & Laguna (1977); it follows that the Gini's coefficient for the κ-generalized distribution is

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

Second, relating the standard deviation to the mean yields the following expression for the Coefficient of Variation

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

Furthermore, the Generalized Entropy class of inequality measures (Cowell, 1980a,b; Cowell & Kuga, 1981a,b) is defined as

$$GE_\kappa(\theta) = \frac{1}{\theta^2-\theta}\left\{ m^{-\theta}\left[\frac{(2\beta\kappa)^{-\frac{\theta}{\alpha}}}{1+\frac{\theta}{\alpha}\kappa} \frac{\Gamma\left(\frac{1}{2\kappa}-\frac{\theta}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{\theta}{2\alpha}\right)} \Gamma\left(1+\frac{\theta}{\alpha}\right)\right] - 1 \right\}, \quad \theta \neq 0, 1. \tag{14}$$

The Mean Logarithmic Deviation is

$$MLD_\kappa = \lim_{\theta\to 0} GE_\kappa(\theta) = \frac{1}{\alpha}\left[\gamma + \psi\left(\frac{1}{2\kappa}\right) + \log(2\beta\kappa) + \alpha\log(m) + \kappa\right],$$

where $\gamma = -\psi(1) = 0.5772156649015328606\ldots$ is the Euler-Mascheroni's constant and $\psi(x) = \mathrm{d}\log\left[\Gamma(x)\right]/\,\mathrm{d}\,x = \Gamma^{'}(x)/\Gamma(x)$ is the Digamma function; the Theil's (1967) index is

$$\begin{aligned} T_\kappa = \lim_{\theta\to 1} GE_\kappa(\theta) =& \frac{(2\beta\kappa)^{-\frac{1}{\alpha}}}{m(\alpha+\kappa)^2} \frac{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right)} \Gamma\left(1+\frac{1}{\alpha}\right) \left\{ (\alpha+\kappa)\left[\psi\left(1+\frac{1}{\alpha}\right) - \right.\right. \\ & \left.\left. -\frac{1}{2}\psi\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right) - \frac{1}{2}\psi\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right) - \alpha\log(m) - \log(2\beta\kappa)\right] - \alpha\kappa \right\}. \end{aligned}$$

Expressions for each index other than for the cases $\theta = 0, 1$ can be derived by straightforward substitution. In particular, the bottom-sensitive index is given by

$$GE_{\kappa}\left(-1\right) = -\frac{1}{2} + \frac{\Gamma\left(1+\frac{1}{\alpha}\right)\Gamma\left(1-\frac{1}{\alpha}\right)}{2\left[1+\left(\frac{\kappa}{\alpha}\right)^{2}\right]},$$

while the expression for the top-sensitive index (or half the squared Coefficient of Variation) is

$$GE_{\kappa}\left(2\right) = \frac{1}{2}\left\{\frac{\frac{\Gamma\left(1+\frac{2}{\alpha}\right)}{1+2\frac{\kappa}{\alpha}}\frac{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{\alpha}\right)}}{\left[\frac{\Gamma\left(1+\frac{1}{\alpha}\right)}{1+\frac{\kappa}{\alpha}}\frac{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right)}\right]^{2}} - 1\right\} = \frac{1}{2}CV_{\kappa}^{2}.$$

Finally, the Atkinson's (1970) index for inequality aversion parameter $\theta = 1-\epsilon$ can be easily computed from $GE_{\kappa}\left(\theta\right)$ by exploiting the relationship $A\left(\epsilon\right) = 1-\left[\epsilon\left(\epsilon-1\right)GE\left(1-\epsilon\right)+1\right]^{\frac{1}{1-\epsilon}}$ (see *e.g.* Cowell, 1995); this yields

$$A_{\kappa}\left(\epsilon\right) = 1-\left\{m^{-(1-\epsilon)}\left[\frac{\left(2\beta\kappa\right)^{-\frac{1-\epsilon}{\alpha}}}{1+\frac{1-\epsilon}{\alpha}\kappa}\frac{\Gamma\left(\frac{1}{2\kappa}-\frac{1-\epsilon}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{1-\epsilon}{2\alpha}\right)}\Gamma\left(1+\frac{1-\epsilon}{\alpha}\right)\right]\right\}^{\frac{1}{1-\epsilon}}, \quad \epsilon \neq 1. \tag{15}$$

The limiting form of Equation (15) as $\epsilon \to 1$ is

$$A_{\kappa}\left(1\right) = 1-e^{-\frac{1}{\alpha}\left[\gamma+\psi\left(\frac{1}{2\kappa}\right)+\log(2\beta\kappa)+\alpha\log(m)+\kappa\right]} = 1-e^{-MLD_{\kappa}}.$$

3.4 Estimation

Parameter estimation for the κ-generalized distribution can be performed using the Maximum Likelihood (ML) approach. Assuming that all observations $\mathbf{x} = \{x_1, \ldots, x_n\}$ are independent, the likelihood function is

$$L\left(\boldsymbol{\theta};\mathbf{x}\right) = \prod_{i=1}^{n} p\left(x_i\right) = \left(\alpha\beta\right)^{n}\prod_{i=1}^{n}\frac{x_i^{\alpha-1}\exp_{\kappa}\left(-\beta x_i^{\alpha}\right)}{\sqrt{1+\beta^2\kappa^2 x_i^{2\alpha}}},$$

where $\boldsymbol{\theta} = \{\alpha, \beta, \kappa\}$ is the parameter vector. This leads to the problem of solving the partial derivatives of the log-likelihood function $l\left(\boldsymbol{\theta};\mathbf{x}\right) = \log L\left(\boldsymbol{\theta};\mathbf{x}\right)$ with respect to κ, α and β. However, obtaining explicit expressions for the ML estimators of the three parameters is difficult, making direct analytical solutions intractable, and one needs to use numerical optimization methods.

Taking into account the meaning of the variable x, the mean value results to be equal to unity, *i.e.* $m = \int_0^{\infty} xp\left(x\right)\mathrm{d}x = 1$. The latter relationship permits to express the parameter β as a function of the parameters κ and α, obtaining

$$\beta = \frac{1}{2\kappa}\left[\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\kappa+\alpha}\frac{\Gamma\left(\frac{1}{2\kappa}-\frac{1}{2\alpha}\right)}{\Gamma\left(\frac{1}{2\kappa}+\frac{1}{2\alpha}\right)}\right]^{\alpha}. \tag{16}$$

In this way, the problem to determine the values of the free parameters $\{\kappa, \alpha, \beta\}$ of the theory from the empirical data reduces to a two parameter $\{\kappa, \alpha\}$ fitting problem. Therefore, to find the parameter values that give the most desirable fit, one can use the Constrained Maximum Likelihood (CML) estimation method (Schoenberg, 1997), which solves the general maximum log-likelihood problem of the form $l(\boldsymbol{\theta};\mathbf{x}) = \sum_{i=1}^{n} \log p(x_i;\boldsymbol{\theta})^{w_i}$, where n is the number of observations, w_i the weight assigned to each observation, $p(x_i;\boldsymbol{\theta})$ the probability of x_i given $\boldsymbol{\theta}$, subject to the non-linear equality constraint given by Equation (16) and bounds $\alpha, \beta > 0$ and $\kappa \in [0,1)$. The CML procedure finds values for the parameters in $\boldsymbol{\theta}$ such that $l(\boldsymbol{\theta};\mathbf{x})$ is maximized using the sequential quadratic programming method (Han, 1977) as implemented, *e.g.*, in MATLAB® 7.

3.5 Empirical Implementation

The κ-generalized distribution was fit to data on personal income distribution for the following countries and years: Germany in 2001, Italy in 2002 and the United Kingdom in 2001. The unit of assessment is the household, and income is expressed in nominal local currency units—and is equivalized for differences in household size by adjusting by the square root of the number of household members (*e.g.* Deaton, 1996). The total number of sample households surveyed in each country is reported in the first row of Table 2. Following the recommendations of The Canberra Group (2001), all calculations use the sampling weights produced by the data provider. The distributions considered are those of disposable income, *i.e.* the income recorded after the payment of taxes and government transfers. Furthermore, observations with zero and negative values have been excluded from the analysis, and income has been normalized to its empirical average (fifth row of Table 1).

The maximum likelihood estimates of the parameters are reported in Table 1. The fit was very good, and the comparison between the predicted probabilities from the present distribution and the observed ones shown in panels (a) and (b) of Figures 7–9 suggests that the κ-generalized functional form offers a great potential for describing the data over their whole range, from the low to medium income region through to the high income Pareto's power-law regime, including the intermediate region for which a clear deviation exists when two different curves are used.

Panel (c) of the same figures depicts the data points for the empirical Lorenz curve, *i.e.* $L\left(\frac{i}{n}\right) = \frac{\sum_{j=1}^{i} x_j}{\sum_{j=1}^{n} x_j}$, $i = 1, 2, \ldots, n$, superimposed by the theoretical curve $L_\kappa(u)$ given by Equation (13) with estimates replacing α and κ as necessary. This formula is shown by the solid line in the plots, and fits the data very well. Furthermore, Table 1 presents the calculated inequality based on the measures discussed in Section 3.3. As it is evident, there is a very good agreement between the estimates of the indexes and the values implied by the κ-generalized distribution[14]; the 95% confidence intervals constructed around the empirical results always cover the theoretical predictions[15].

In order to further evaluate the accuracy of the proposed distributional model, the hypothesis that each set of n observed data follows a κ-generalized distribution has also been tested by calculating the K-S statistic $D^{+} = \max_{1\leq i\leq n}\left[in^{-1} - P_{\leq}(x_i)\right]$, $i = 1, 2, \ldots, n$. Since in this case there is no asymptotic formula for calculating the p-value, the problem has been reduced to testing that the x values have a standard exponential distribution (*i.e.*, an exponential distribution with parameter equal to 1) by relating the function $P_{>}(x)$ given by Equation (10) to the ordinary exponential function, namely $\exp_\kappa(-\beta x^\alpha) = \exp(-x_\kappa)$, through the transformation $x_\kappa = \frac{1}{\kappa}\log\left(\sqrt{1+\beta^2\kappa^2x^{2\alpha}} + \beta\kappa x^\alpha\right)$, where the parameters are estimated from the data. Thus

[14] For the formulas used to estimate the inequality measures of Section 3.3 see *e.g.* Cowell (1995).

[15] The upper and lower confidence limits have been obtained via the "bootstrap-t" method, which both theory and Monte Carlo evidence have shown to provide better coverage in many aplications (*e.g.* Hall, 1988, 1992, and Horowitz, 2001). To have an available estimate of the standard error of the parameter of interest for each bootstrap replication, a bootstrap-within-bootstrap procedure based on a total number of replications equal to $B_1 \times B_2 = 100 \times 25 = 2500$—where B_2 is the number of resamples drawn to obtain an estimate of the standard error used to compute the sequence $\{t_i\}_{i=1}^{i=B_1}$—has been used.

Table 2: *Estimated parameters of the κ-generalized distribution for the countries and years shown in Figures 7–9. Also shown are the total number of sample households surveyed, the estimated weighted average income and corresponding 95% confidence interval, the empirical estimates and theoretical predictions of the inequality measures discussed in Section 3.3, and the value of the K-S goodness-of-fit test statistic.*

	Germany	**Italy**	**United Kingdom**
$\boldsymbol{n}$	11,344	8,011	10,636
$\boldsymbol{\alpha}^{a}$	2.5659 (2.5653, 2.5666)	2.2540 (2.2533, 2.2547)	2.7357 (2.7348, 2.7366)
$\boldsymbol{\beta}^{a}$	0.8788 (0.8786, 0.8791)	1.0087 (1.0083, 1.0091)	0.9433 (0.9429, 0.9437)
$\boldsymbol{\kappa}^{a}$	0.5697 (0.5692, 0.5702)	0.6944 (0.6937, 0.6950)	0.7080 (0.7074, 0.7086)
$\langle \boldsymbol{z} \rangle$	36315.67 (35976.42, 36654.91)	18087.92 (17841.07, 18334.77)	14982.20 (14799.11,15165.28)
$\boldsymbol{G}^{b}$	0.2749 (0.2679, 0.2803)	0.3311 (0.3261, 0.3358)	0.2790 (0.2735, 0.2891)
$\boldsymbol{G_{\kappa}}$	0.2748	0.3306	0.2772
$\boldsymbol{CV}^{b}$	0.5505 (0.4918, 0.6094)	0.7344 (0.6937, 0.7842)	0.6073 (0.5171, 0.8136)
$\boldsymbol{CV_{\kappa}}$	0.5368	0.7366	0.5712
$\boldsymbol{MLD}^{b}$	0.1430 (0.1338, 0.1529)	0.2015 (0.1948, 0.2089)	0.1525 (0.1345, 0.1612)
$\boldsymbol{MLD_{\kappa}}$	0.1362	0.1971	0.1360
$\boldsymbol{T}^{b}$	0.1306 (0.1176, 0.1411)	0.1960 (0.1873, 0.2053)	0.1402 (0.1259, 0.1616)
$\boldsymbol{T_{\kappa}}$	0.1276	0.1948	0.1334
$\boldsymbol{GE\,(2)}^{b}$	0.1515 (0.1171, 0.1858)	0.2697 (0.2378, 0.3086)	0.1844 (0.1297, 0.3110)
$\boldsymbol{GE_{\kappa}\,(2)}$	0.1441	0.2713	0.1631
$\boldsymbol{A\,(1)}^{b}$	0.1332 (0.1254, 0.1416)	0.1825 (0.1771, 0.1885)	0.1415 (0.1256, 0.1491)
$\boldsymbol{A_{\kappa}\,(1)}$	0.1274	0.1789	0.1272
$\boldsymbol{D^{+}}^{c}$	0.0084 (0.1999)	0.0069 (0.1331)	0.0091 (0.2306)

[a] In parentheses: normal-approximation 95% confidence interval based on the reciprocal of the observed Fisher's information matrix.

[b] In parentheses: bootstrap-t 95% confidence interval based on 2500 replications.

[c] In parentheses: p-value.

Source: author's own calculations based on the GSOEP-CNEF income data for the year 2001, the SHIW income data for the year 2002 and the BHPS-CNEF income data for the year 2001.

the significance level in the upper tail is given approximatively by $P_{>}(T^{*}) = \exp\left[-2\left(T^{*}\right)^{2}\right]$, with $T^{*} = D^{+}\left(\sqrt{n} + 0.12 + 0.11/\sqrt{n}\right)$, as suggested for example by Stephens (1970). The results are shown in the last row of Table 1. As one can appreciate, the maximum distance between the empirical data and the theoretical model as assessed by the K-S statistic is very small, and the p-values in parentheses do not lead to rejection of the null hypothesis that the data may come from a κ-generalized distribution at any of the usual significance levels (1%, 5% and 10%). The linear behaviour emerging from the Quantile-Quantile (Q-Q) plots of the sample quantiles versus the corresponding quantiles of the fitted κ-generalized distribution displayed in panel (d) of Figs. 7–9 strongly supports the quantitative results obtained by hypothesis testing.

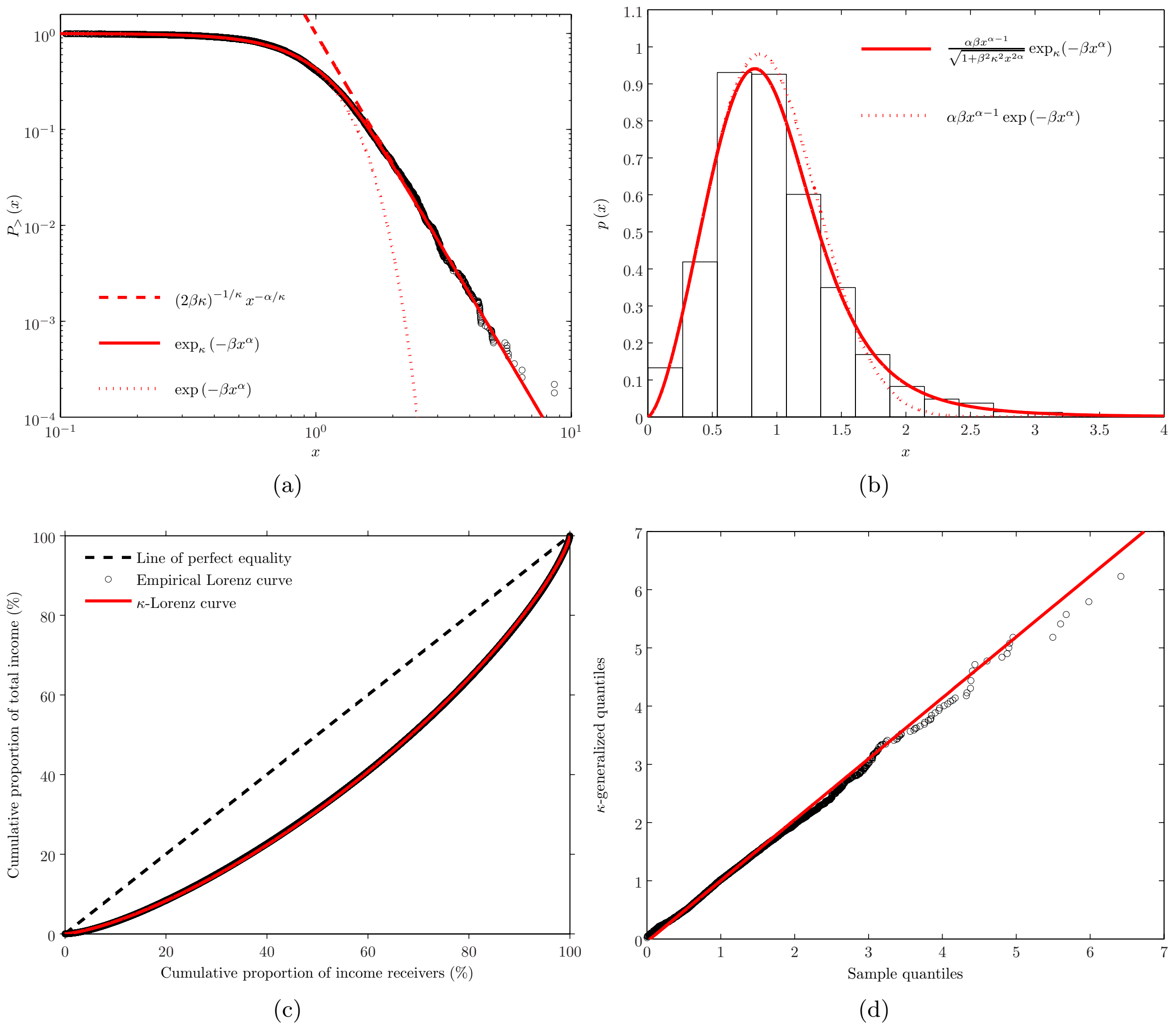


Figure 7: *The German personal income distribution in 2001. The income variable is measured in current year euros. (a) Plot of the empirical CCDF in the log-log scale. The solid line is the theoretical model given by Equation* (10) *fitting very well the data in the whole range from the low to the high incomes including the intermediate income region. This function is compared with the ordinary stretched exponential one (dotted line)—fitting the low income data—and with the pure power-law (dashed line)—fitting the high income data. (b) Histogram plot of the empirical PDF with superimposed fits of the κ-generalized (solid line) and Weibull (dotted line) PDFs. (c) Plot of the Lorenz curve. The hollow circles represent the empirical data points and the solid line is the theoretical curve given by Equation* (13)*. The dashed line corresponds to the Lorenz curve of a society in which everybody receives the same income and thus serves as a benchmark case against which actual income distribution may be measured. (d) Q-Q plot of the sample quantiles versus the corresponding quantiles of the fitted κ-generalized distribution. The reference line has been obtained by locating points on the plot corresponding to around the 25th and 75th percentiles and connecting these two. In plots (a), (b) and (d) the income axis limits have been adjusted according to the range of data to shed light on the intermediate region between the bulk and the tail of the distribution.*

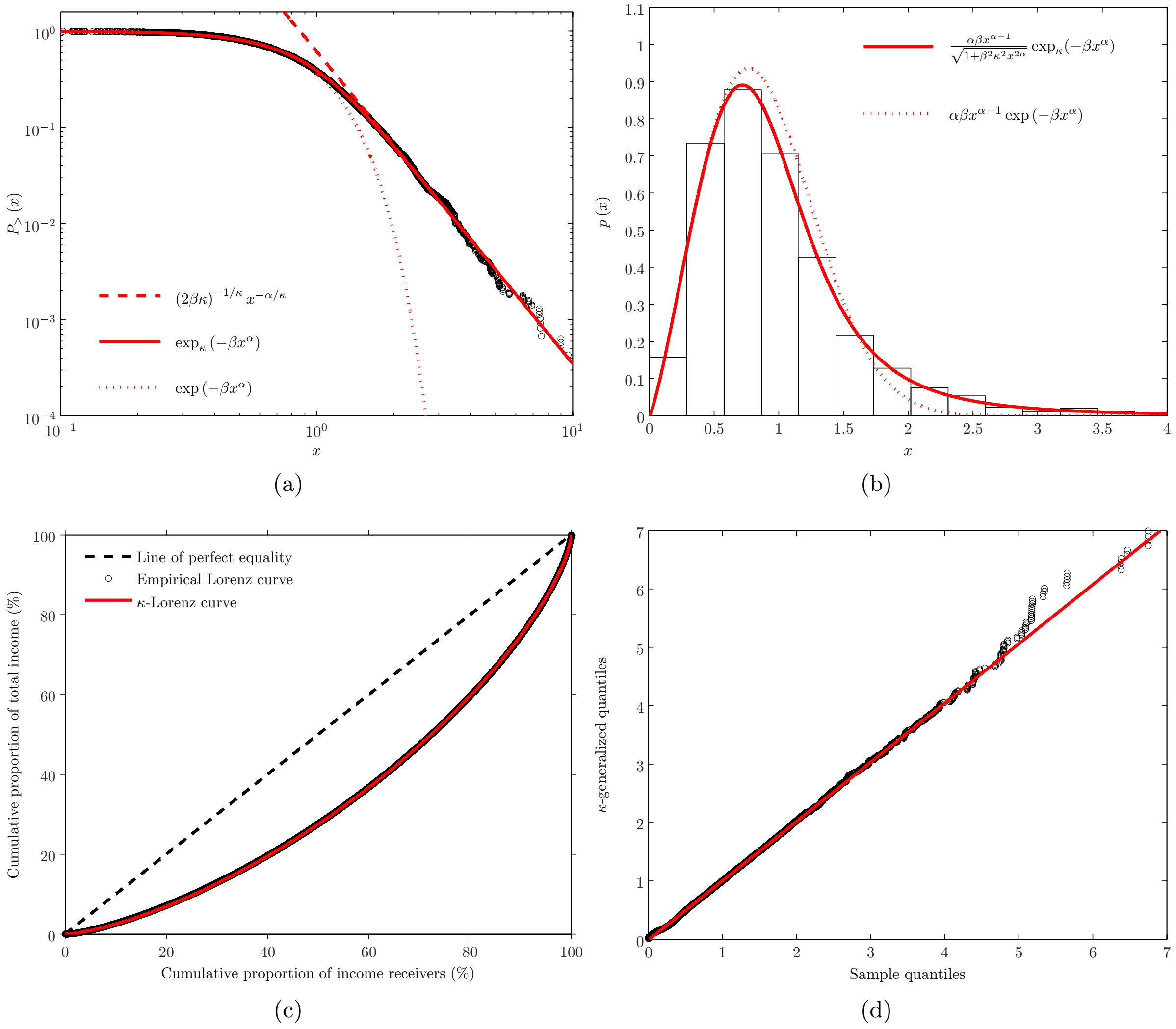


Figure 8: *Same plots as in Figure 7 for the Italian personal income distribution in 2002. The income variable is measured in current year euros.*

References

Aigner, D. J., & Goldberger, A. S. (1970). Estimation of Pareto's law from grouped observations. *Journal of the American Statistical Association*, *65*, 712–723.

Aitchison, J., & Brown, J. A. C. (1954). On criteria for descriptions of income distribution. *Metroeconomica*, *6*, 81–144.

Aitchison, J., & Brown, J. A. C. (1957). *The Lognormal Distribution with Special Reference to its Use in Economics.* New York: Cambridge University Press.

Andrews, D. W. K., & Buchinsky, M. (2000). A three-step method for choosing the number of bootstrap replications. *Econometrica*, *68*, 23–51.

Arnold, B. C. (1983). *Pareto Distributions.* Fairland: International Co-operative Publishing House.

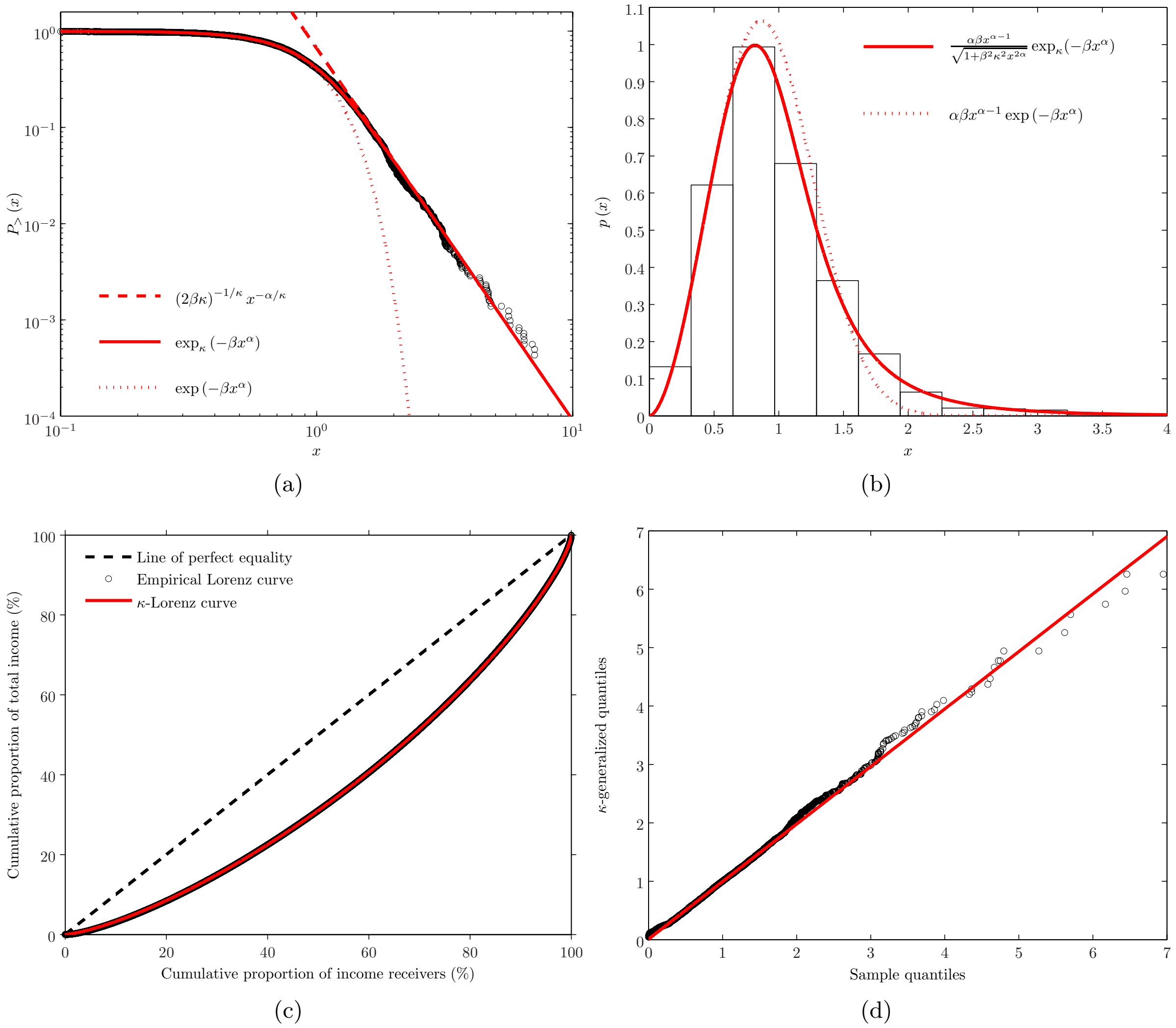


Figure 9: *Same plots as in Figures 7 and 8 for the UK personal income distribution in 2001. The income variable is measured in current British pounds.*

Arnold, B. C., & Laguna, L. (1977). *On Generalized Pareto Distributions with Applications to Income Data*. Ames: Iowa State University Press.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*, 244–263.

Atkinson, A. B., & Harrison, A. J. (1978). *The Distribution of Wealth in Britain*. Cambridge: Cambridge University Press.

Atoda, N., Suruga, T., & Tachibanaki, T. (1988). Statistical inference of functional forms for income distribution. *The Economic Studies Quarterly*, *39*, 14-40.

Bartels, C. P. A. (1977). *Economic Aspects of Regional Welfare: Income Distribution and Unemployment*. Leiden: Martinus Nijhoff.

Bordley, R. F., McDonald, J. B., & Mantrala, A. (1996). Something new, something old: Parametric models for the size distribution of income. *Journal of Income Distribution*, *6*, 91-103.

Brachmann, K., Stich, A., & Trede, M. (1996). Evaluating parametric income distribution models. *Allgemeines Statistisches Archiv*, *80*, 285-298.

Champernowne, D. G. (1953). A model of income distribution. *The Economic Journal*, *63*, 318–351.

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2007). *Power-Law Distributions in Empirical Data* (SFI Working Paper No. 07-12-049). Available from `http://www.santafe.edu/research/publications/wpabstract/200712049`

Coronel-Brizio, H. F., & Hernández-Montoya, A. R. (2005). On fitting the Pareto-Levy distribution to stock market index data: Selecting a suitable cutoff value. *Physica A: Statistical Mechanics and its Applications*, *354*, 437–449.

Cowell, F. A. (1980a). Generalized entropy and the measurement of distributional change. *European Economic Review*, *13*, 147–159.

Cowell, F. A. (1980b). On the structure of additive inequality measures. *Review of Economic Studies*, *47*, 521–531.

Cowell, F. A. (1995). *Measuring Inequality.* Hemel Hempstead: Harvester Wheatsheaf.

Cowell, F. A., & Kuga, K. (1981a). Additivity and the entropy concept: An axiomatic approach to inequality measure. *Journal of Economic Theory*, *25*, 131–143.

Cowell, F. A., & Kuga, K. (1981b). Inequality measurement: An axiomatic approach. *European Economic Review*, *15*, 287–305.

Dacorogna, M. M., Müller, U. A., Pictet, O. V., & De Vries, C. G. (1992). *The Distribution of Extremal Foreign Exchange Rate Returns in Extremely Large Data Sets* (Working Paper No. 1992–10–22). Zürich: Olsen & Associates. Available from `http://ideas.repec.org/p/wop/olaswp/_012.html`

D'Agostino, R. B., & Stephens, M. A. (Eds.). (1986). *Goodness-of-Fit Techniques.* New York: Marcel Dekker.

Dagum, C. (1977). A new model of personal income distribution: Specification and estimation. *Economie Appliquée*, *30*, 413–436.

Dalton, H. (1920). The measurement of the inequality of incomes. *The Economic Journal*, *30*, 348–361.

Danielsson, J., De Haan, L., Peng, L., & De Vries, C. G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*, *76*, 226–248.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and their Application.* New York: Cambridge University Press.

Deaton, A. (1996). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Baltimore MD: Johns Hopkins University Press.

Efron, B., & Tibshirani, R. (1993). *An Introduction to the Bootstrap.* New York: Chapman & Hall.

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Berlin: Springer-Verlag.

Espinguet, P., & Terraza, M. (1983). Essai d'extrapolation des distributions de salaires français. *Economie Appliquée*, *36*, 535–561.

Gastwirth, J. L. (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.* Paris: Librairie du Recueil Sirey.

Gini, C. (1914). Sulla misura della concentrazione e della variabilità dei caratteri. In *Transactions of the Real Istituto Veneto di Scienze, Lettere ed Arti* (Vol. LIII, pp. 1203–1248). Venice: Premiate Officine Grafiche C. Ferrari. (English translation in *Metron - International Journal of Statistics*, *63*, 1–38, 2005)

Goldstein, M. L., Morris, S. A., & Yen, G. G. (2004). Problems with fitting to the power-law distribution. *European Physical Journal B*, *41*, 255–258.

Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society. Series B (Methodological)*, *44*, 37–42.

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. *The Annals of Statistics*, *16*, 927–953.

Hall, P. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis*, *32*, 177–203.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* New York: Springer-Verlag.

Hall, P., & Welsh, A. H. (1985). Adaptive estimates of parameters of regular variation. *The Annals of Statistics*, *13*, 331–341.

Han, S.-P. (1977). A globally convergent method for nonlinear programming. *Journal of Optimization Theory and Applications*, *22*, 297–309.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, *3*, 1163–1174.

Horowitz, J. L. (2001). The bootstrap. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of Econometrics* (Vol. 5). Amsterdam: North-Holland.

Jenkins, S. P., & Jäntti, M. (2005). *Methods for Summarizing and Comparing Wealth Distributions* (ISER working papers No. 2005-05). Colchester: Institute for Social and Economic Research. Available from `http://ideas.repec.org/p/ese/iserwp/2005-05.html`

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous Univariate Distributions* (2nd ed., Vol. 1). New York: John Wiley & Sons.

Kaniadakis, G. (2001). Non-linear kinetics underlying generalized statistics. *Physica A: Statistical Mechanics and its Applications*, *296*, 405-425.

Kaniadakis, G. (2002). Statistical mechanics in the context of special relativity. *Physical Review E*, *66*, 056125.

Kaniadakis, G. (2005). Statistical mechanics in the context of special relativity. II. *Physical Review E*, *72*, 036108.

Kleiber, C., & Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences.* New York: John Wiley & Sons.

Laherrère, J., & Sornette, D. (1998). Stretched exponential distributions in nature and economy: "Fat tails" with characteristic scales. *The Europen Physical Journal B*, *2*, 525–539.

Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, *9*, 209–219.

Lux, T. (2000). On moment condition failure in German stock returns: An application of recent advances in extreme value statistics. *Empirical Economics*, *25*, 641–652.

Lux, T. (2001). The limiting extremal behaviour of speculative returns: An analysis of intradaily data from the Frankfurt Stock Exchange. *Applied Financial Economics*, *11*, 299–315.

Majumder, A., & Chakravarty, S. R. (1990). Distribution of personal income: Development of a new model and its application to U. S. income data. *Journal of Applied Econometrics*, *5*, 189–196.

Mandelbrot, B. (1960). The Pareto-Levy law and the distribution of income. *International Economic Review*, *1*, 79–106.

Mason, D. M. (1982). Laws of large numbers for sums of extreme values. *The Annals of Probability*, *10*, 754–764.

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, *52*, 647–665.

McDonald, J. B., & Ransom, M. R. (1979). Functional forms, estimation techniques and the distribution of income. *Econometrica*, *47*, 1513–1525.

Metcalf, C. E. (1969). The size distribution of personal income during the business cycle. *The American Economic Review*, *59*, 657–668.

Montroll, E. W., & Shlesinger, M. F. (1982). On $1/f$ noise and other distributions with long tails. *Proceedings of the National Academy of Sciences USA*, *79*, 3380–3383.

Montroll, E. W., & Shlesinger, M. F. (1983). Maximum entropy formalism, fractals, scaling phenomena, and $1/f$ noise: A tale of tails. *Journal of Statistical Physics*, *32*, 209–230.

Ord, J. K., Patil, G. P., & Taillie, C. (1981). The choice of a distribution to describe personal incomes. In C. Taillie, G. P. Patil, & B. A. Baldessari (Eds.), *Statistical Distributions in Scientific Work* (Vol. 6, pp. 193–201). Dordrecht: D. Reidel Publishing Company.

Organisation for Economic Co-operation and Development. (2003). *Statistical Compendium* (ed. 02#2003).

Pareto, V. (1964). Course d'économie politique. In G.-H. Bousquet & G. Busino (Eds.), *Œuvres complètes de Vilfredo Pareto, Tome 1.* Geneva: Libraire Droz.

Pareto, V. (1965). La courbe de la repartition de la richesse. In G. Busino (Ed.), *Œuvres complètes de Vilfredo Pareto, Tome 3: Écrits sur la courbe de la répartition de la richesse* (pp. 1–15). Geneva: Librairie Droz. (English translation in *Rivista di Politica Economica*, *87*, 647–700, 1997)

Persky, J. (1992). Retrospectives: Pareto's law. *The Journal of Economic Perspectives*, *6*, 181–192.

Pictet, O. V., Dacorogna, M. M., & Müller, U. A. (1996). *Hill, Bootstrap and Jackknife Estimators for Heavy Tails* (Working Paper No. 1996–12–10). Zürich: Olsen & Associates. Available from `http://ideas.repec.org/p/wop/olaswp/_015.html`

Quandt, R. E. (1996). Old and new methods of estimation and the Pareto distribution. *Metrika*, *10*, 55–82.

Rajaonarison, D., Bolduc, D., & Jayet, H. (2005). The K-deformed multinomial logit model. *Economics Letters*, *86*, 13-20.

Reed, W. J., & Jorgensen, M. (2004). The double Pareto-lognormal distribution – A new parametric model for size distribution. *Communications in Statistics – Theory and Methods*, *33*, 1733-1753.

Resnick, S. I. (1997). Heavy tail modeling in teletraffic data. *The Annals of Statistics*, *25*, 1866–1869.

Salem, A. B. Z., & Mount, T. D. (1974). A convenient descriptive model of income distribution: The gamma density. *Econometrica*, *42*, 1115–1127.

Schoenberg, R. J. (1997). Constrained maximum likelihood. *Computational Economics*, *10*, 251–266.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423 and 623–657.

Singh, S. K., & Maddala, G. S. (1976). A function for size distribution of incomes. *Econometrica*, *44*, 963–970.

Sornette, D. (2004). *Critical Phenomena in Natural Sciences. Chaos, Fractals, Selforganization and Disorder: Concepts and Tools.* Berlin: Springer-Verlag.

Steindl, J. (1965). *Random Processes and the Growth of Firms: A Study of the Pareto Law.* New York: Hafner Press.

Stephens, M. A. (1970). Use of the Kolmogorov-Smirnov, Cramér-Von Mises and related statistics without extensive tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *32*, 115–122.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, *69*, 730–737.

Tachibanaki, T., Suruga, T., & Atoda, N. (1997). Estimations of income distribution parameters for individual observations by maximum likelihood method. *Journal of the Japan Statistical Society*, *27*, 191-203.

The Canberra Group. (2001). *Final Report and Recommendations.* Available from `http://www.lisproject.org/links/canberra/finalreport.pdf`

Theil, H. (1967). *Economics and Information Theory.* Amsterdam: North-Holland.

Weron, R. (2001). Levy-stable distributions revisited: Tail index $\alpha > 2$ does not exclude the levy-stable regime. *International Journal of Modern Physics C*, *12*, 209–223.